

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OK protein - protein search, using sw model

Run on: November 13, 2002, 13:40:23 ; Search time 14 Seconds
(without alignments)
8.407 Million cell updates/sec

Title: US-09-697-590-2_COPY_292_295

Perfect score: 21

Sequence: 1 HARL 4

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Issued_Patents_AA: *
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep: *
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep: *
3: /cgn2_6/ptodata/1/1aa/5A.COMB.pep: *
4: /cgn2_6/ptodata/1/1aa/5B.COMB.pep: *
5: /cgn2_6/ptodata/1/1aa/PCITUS.COMB.pep: *
6: /cgn2_6/ptodata/1/1aa/Backfilest1.pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	21	100.0	15	4	US-09-347-504-66 Sequence 66, Appl
2	21	100.0	17	1	US-08-399-646-8 Sequence 8, Appl
3	21	100.0	17	1	US-08-607-321-8 Sequence 8, Appl
4	21	100.0	17	2	US-08-961-240-8 Sequence 8, Appl
5	21	100.0	17	2	US-08-605-501-8 Sequence 8, Appl
6	21	100.0	51	2	US-08-870-518-23 Sequence 23, Appl
7	21	100.0	53	1	US-08-222-616-6 Sequence 6, Appl
8	21	100.0	53	4	US-08-446-648-6 Sequence 6, Appl
9	21	100.0	53	5	PCIT-US95-04228-6 Sequence 6, Appl
10	21	100.0	82	3	US-08-403-852D-27 Sequence 27, Appl
11	21	100.0	82	3	US-08-510-646B-28 Sequence 28, Appl
12	21	100.0	82	4	US-09-231-818-27 Sequence 27, Appl
13	21	100.0	91	2	US-09-047-125-10 Sequence 10, Appl
14	21	100.0	91	3	US-07-736-335E-10 Sequence 10, Appl
15	21	100.0	113	4	US-09-311-352B-2 Sequence 2, Appl
16	21	100.0	120	1	US-08-420-235B-29 Sequence 29, Appl
17	21	100.0	120	4	US-08-793-624-29 Sequence 29, Appl
18	21	100.0	120	4	PCIT-US95-10194-29 Sequence 29, Appl
19	21	100.0	166	4	US-09-134-001C-3657 Sequence 3657, Ap
20	21	100.0	170	3	US-09-186-250-4 Sequence 4, Appl
21	21	100.0	170	3	US-09-186-250-8 Sequence 4, Appl
22	21	100.0	170	4	US-09-517-347-4 Sequence 4, Appl
23	21	100.0	170	4	US-09-517-347-8 Sequence 4, Appl
24	21	100.0	170	4	US-09-518-232A-4 Sequence 4, Appl
25	21	100.0	170	4	US-09-518-232A-8 Sequence 4, Appl
26	21	100.0	170	4	US-09-517-358A-4 Sequence 4, Appl
27	21	100.0	170	4	US-09-517-358A-8 Sequence 8, Appl

ALIGNMENTS

28	21	100.0	179	2	US-08-469-537A-84	Sequence 84, Appl
29	21	100.0	183	4	US-09-311-352B-4	Sequence 4, Appl
30	21	100.0	192	2	US-08-469-537A-37	Sequence 37, Appl
31	21	100.0	208	1	US-07-661-610C-5	Sequence 5, Appl
32	21	100.0	208	1	US-08-309-512-4	Sequence 4, Appl
33	21	100.0	208	5	PCIT-US92-08756A-4	Sequence 4, Appl
34	21	100.0	209	4	US-09-199-637A-401	Sequence 401, App
35	21	100.0	213	3	US-09-186-250-2	Sequence 2, Appl
36	21	100.0	213	3	US-09-186-250-6	Sequence 6, Appl
37	21	100.0	213	4	US-09-517-347-2	Sequence 2, Appl
38	21	100.0	213	4	US-09-517-347-6	Sequence 6, Appl
39	21	100.0	213	4	US-09-518-232A-2	Sequence 2, Appl
40	21	100.0	213	4	US-09-518-232A-6	Sequence 6, Appl
41	21	100.0	213	4	US-09-517-358A-2	Sequence 2, Appl
42	21	100.0	213	4	US-09-517-358A-6	Sequence 6, Appl
43	21	100.0	215	2	US-08-796-676-3	Sequence 3, Appl
44	21	100.0	215	2	US-08-967-364-8	Sequence 8, Appl
45	21	100.0	215	3	US-09-368-408-8	Sequence 8, Appl

RESULT 1
US-09-347-504-66
Sequence 66, Application US/09347504
Patent No. 6399075
GENERAL INFORMATION:
APPLICANT: Howley, Peter M.
APPLICANT: Benson, John
APPLICANT: Kasukawa, Hiroaki
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
FILE REFERENCE: HMV-041.01
CURRENT APPLICATION NUMBER: US/09/347,504
CURRENT FILING DATE: 1999-07-02
NUMBER OF SEQ ID NOS: 79
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 66
LENGTH: 15
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: M36A
US-09-347-504-66

Query Match 100.0%; Score 21; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HARL 4
Db 3 HARL 6
RESULT 2
US-08-399-646-8
Sequence 8, Application US/08399646
Patent No. 5556781
GENERAL INFORMATION:
APPLICANT: KUBOTA, Michio
APPLICANT: TSUSAKI, Keiji
APPLICANT: HATTORI, Kazuko
APPLICANT: SUGIMOTO, Toshiyuki
TITLE OF INVENTION: DNA ENCODING ENZYME, RECOMBINANT DNA AND
NUMBER OF INVENTION: ENZYME, TRANSFORMANT, AND THEIR PREPARATIONS AND USE
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESSES:
ADDRESSEE: BROWDY AND NEIMARK
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA

ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/399,646
FILING DATE: 07-MAR-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 59834
FILING DATE: 07-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 59840
FILING DATE: 07-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: BROWDY, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: KUBOTA-5
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 17 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-399-646-8

Query Match 100.0%; Score 21; DB 1; Length 17;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
||||
DB 11 HARL 14

RESULT 3
US-08-607-321-8
Sequence 8, Application US/08607321
Patent No. 5718613
GENERAL INFORMATION:
APPLICANT: KUBOTA, Michio
APPLICANT: TSUSAKI, Keiji
APPLICANT: HATTORI, Kazuko
APPLICANT: SUGIMOTO, Toshiyuki
TITLE OF INVENTION: DNA ENCODING ENZYME, RECOMBINANT DNA AND
TITLE OF INVENTION: ENZYME, TRANSFORMANT, AND THEIR PREPARATIONS AND USE
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/607,321
FILING DATE: 26-FEB-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/399,646
FILING DATE: 07-MAR-1995

APPLICATION NUMBER: JP 59834
FILING DATE: 07-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 59840
FILING DATE: 07-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: BROWDY, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: KUBOTA-5
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 17 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-607-321-8

Query Match 100.0%; Score 21; DB 1; Length 17;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
||||
DB 11 HARL 14

RESULT 4
US-08-961-240-8
Sequence 8, Application US/08961240
Patent No. 5830715
GENERAL INFORMATION:
APPLICANT: KUBOTA, Michio
APPLICANT: TSUSAKI, Keiji
APPLICANT: HATTORI, Kazuko
APPLICANT: SUGIMOTO, Toshiyuki
TITLE OF INVENTION: DNA ENCODING ENZYME, RECOMBINANT DNA AND
TITLE OF INVENTION: ENZYME, TRANSFORMANT, AND THEIR PREPARATIONS AND USE
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,240
FILING DATE: 30-OCT-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/399,646
FILING DATE: 07-MAR-1995
APPLICATION NUMBER: JP 59834
FILING DATE: 07-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 59840
FILING DATE: 07-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: BROWDY, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: KUBOTA-5
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197

TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 17 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-961-240-8

Query Match
Best Local Similarity 100.0%; Score 21; DB 2; Length 17;
Pred. No. 25;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
1111
DB 11 HARL 14

RESULT 5
US-08-605-501-8
Sequence 8, Application US/08605501
Patent No. 5834287
GENERAL INFORMATION:
APPLICANT: KUBOTA, Michio
APPLICANT: TSUSAKI, Keiji
APPLICANT: HATTORI, Kazuko
APPLICANT: SUGIMOTO, Toshiyuki
TITLE OF INVENTION: DNA ENCODING ENZYME, RECOMBINANT DNA AND
TITLE OF INVENTION: ENZYME, TRANSFORMANT, AND THEIR PREPARATIONS AND USE
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESSES:
ADDRESSEE: BROWDY AND NEIMARK
STREET: 419 Seventh Street, N.W., Suite 300
City: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/605,501
FILING DATE: 26-FEB-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/399,646
FILING DATE: 07-MAR-1995
APPLICATION NUMBER: JP 59834
FILING DATE: 07-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 59840
FILING DATE: 07-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: BROWDY, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: KUBOTA-5
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 17 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-605-501-8

Query Match
Best Local Similarity 100.0%; Score 21; DB 2; Length 17;
Pred. No. 25;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
1111
DB 11 HARL 14

RESULT 6
US-08-870-518-23
Sequence 23, Application US/08870518
Patent No. 5925566
GENERAL INFORMATION:
APPLICANT: Davis, Roger J.
APPLICANT: Galcheva-Gargova, Zoya
TITLE OF INVENTION: NON-ACTIVATED RECEPTOR COMPLEX
TITLE OF INVENTION: PROTEINS AND USES THEREOF
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
City: Boston
STATE: MA
COUNTRY: US
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/870,518
FILING DATE: 06-JUN-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/019,219
FILING DATE: 06-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Fasse, Peter J.
REGISTRATION NUMBER: 32,983
REFERENCE/DOCKET NUMBER: 04020/102001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 51 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-870-518-23

Query Match
Best Local Similarity 100.0%; Score 21; DB 2; Length 51;
Pred. No. 75;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
1111
DB 48 HARL 51

RESULT 7
US-08-222-616-6
Sequence 6, Application US/08222616
Patent No. 5635177
GENERAL INFORMATION:
APPLICANT: Bennett, Brian D.
APPLICANT: Goeddel, David
APPLICANT: Lee, James M.
APPLICANT: Mathews, William
APPLICANT: Tsai, Siao Ping
APPLICANT: Wood, William I.

TITLE OF INVENTION: PROTEIN TYROSINE KINASE AGONIST
NUMBER OF INVENTIONS: ANTIBODIES
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/222,616
FILING DATE: 4-APR-1994
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/00586
FILING DATE: 22-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/826935
FILING DATE: 22-JAN-1992
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M.
REGISTRATION NUMBER:
REFERENCE/DOCKET NUMBER: 821P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 53 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-222-616-6

Query Match 100.0%; Score 21; DB 1; Length 53;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HARL 4
Db 41 HARL 44

RESULT 8
US-08-446-648-6
Sequence 6, Application US/08446648
Patent No. 6331302
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
APPLICANT: Bennett, Brian D.
APPLICANT: Goeddel, David
APPLICANT: Lee, James M.
APPLICANT: Matthews, William
APPLICANT: Tsai, Siao Ping
APPLICANT: Wood, William I.
TITLE OF INVENTION: PROTEIN TYROSINE KINASE AGONIST ANTIBODIES
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/446,648
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/222616
FILING DATE: 04-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M.
REGISTRATION NUMBER: 40,378
REFERENCE/DOCKET NUMBER: P0821P3PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 53 amino acids
TYPE: Amino Acid
TOPOLOGY: linear
US-08-446-648-6

Query Match 100.0%; Score 21; DB 4; Length 53;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HARL 4
Db 41 HARL 44

RESULT 9
PCT-US95-04228-6
Sequence 6, Application PC/TUS9504228
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
APPLICANT: Bennett, Brian D.
APPLICANT: Goeddel, David
APPLICANT: Lee, James M.
APPLICANT: Matthews, William
APPLICANT: Tsai, Siao Ping
APPLICANT: Wood, William I.
TITLE OF INVENTION: PROTEIN TYROSINE KINASE AGONIST ANTIBODIES
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/04228
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/222616
FILING DATE: 04-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Wendy M. Lee
REGISTRATION NUMBER: 00,000
REFERENCE/DOCKET NUMBER: 821P3PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
TELEX: 910/371-7168

INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 53 amino acids
TYPE: amino acid
TOPOLOGY: linear
PCT-US95-04228-6

Query Match 100.0%; Score 21; DB 5; Length 53;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HARL 4
DB 41 HARL 44

RESULT 10
US-08-403-852D-27
Sequence 27, Application US/08403852D
Patent No. 5891695

GENERAL INFORMATION:
APPLICANT: Blanc, Veronique
APPLICANT: Blanc, Francis
APPLICANT: Crouzet, Joel
APPLICANT: Jacques, Nathalie
APPLICANT: Lacroix, Patricia
APPLICANT: Thibaut, Denis
APPLICANT: Zagorec, Monique
APPLICANT: Debussche, Laurent
TITLE OF INVENTION: Polypeptides Involved In The
TITLE OF INVENTION: Biosynthesis Of Streptogramins, Nucleotide Sequences
TITLE OF INVENTION: Coding For These Polypeptides And Their Use
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner
STREET: 1300 I Street, N.W., Suite 700
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/403,852D
FILING DATE: 10-MAY-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR 93/00923
FILING DATE: 25-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 92/11441
FILING DATE: 25-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Kenneth J.
REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 03806, 0054-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 408-4000
TELEFAX: (202) 408-4400
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 82 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

US-08-403-852D-27
Query Match 100.0%; Score 21; DB 2; Length 82;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HARL 4
DB 46 HARL 49

RESULT 11
US-08-510-646B-28
Sequence 28, Application US/08510646B
Patent No. 6077699

GENERAL INFORMATION:
APPLICANT: Blanc, Veronique
APPLICANT: Blanc, Francis
APPLICANT: Crouzet, Joel
APPLICANT: Jacques, Nathalie
APPLICANT: Lacroix, Patricia
APPLICANT: Thibaut, Denis
APPLICANT: Zagorec, Monique
APPLICANT: Debussche, Laurent
TITLE OF INVENTION: Polypeptides Involved In The
TITLE OF INVENTION: Biosynthesis Of Streptogramins, Nucleotide Sequences
TITLE OF INVENTION: Coding For These Polypeptides And Their Use
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner
STREET: 1300 I Street, N.W., Suite 700
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/510,646B
FILING DATE: 03-AUG-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/403,852
FILING DATE: 10-MAY-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR 93/00923
FILING DATE: 25-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 92/11441
FILING DATE: 25-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Kenneth J.
REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 03806, 0054-01000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 408-4000
TELEFAX: (202) 408-4400
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 82 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

US-08-510-646B-28
Query Match 100.0%; Score 21; DB 3; Length 82;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HARL 4
DB 46 HARL 49

RESULT 12
US-09-231-818-27
Sequence 27, Application US/09231818
Patent No. 6171846
GENERAL INFORMATION:
APPLICANT: Blanc, Veronique
APPLICANT: Blanche, Francis
APPLICANT: Crouzet, Joel
APPLICANT: Jacques, Nathalie
APPLICANT: Lacroix, Patricia
APPLICANT: Thibault, Denis
APPLICANT: Zagorec, Monique
APPLICANT: Debussche, Laurent
TITLE OF INVENTION: Polypeptides Involved In The
Biosynthesis Of Streptogramins, Nucleotide Sequences
TITLE OF INVENTION: Coding For These Polypeptides And Their Use
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finegan, Henderson, Farbow, Garrett & Dunner
STREET: 1300 I Street, N.W., Suite 700
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/231.818
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/403.852
FILING DATE: 10-MAY-1995
APPLICATION NUMBER: PCT/FR 93/00923
FILING DATE: 25-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 92/11441
FILING DATE: 25-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Kenneth J.
REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 03806.0054-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 408-4000
TELEFAX: (202) 408-4400
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 82 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-231-818-27
Query Match 100.0%; Score 21; DB 4; Length 82;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 13
US-09-047-125-10
Sequence 10, Application US/09047125
Patent No. 5976787
GENERAL INFORMATION:
APPLICANT: Leland F. Velicer, Peter Brunovskis,
TITLE OF INVENTION: Marek's Disease Herpesvirus
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ian C. McLeod
STREET: 2190 Commons Parkway
CITY: Okemos
STATE: Michigan
COUNTRY: USA
ZIP: 48864
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 1.2 Mb
COMPUTER: IBM PS2, Model 50
OPERATING SYSTEM: MS-DOS 5.0
SOFTWARE: PC-Write 3.02
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/047.125
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/736.335
FILING DATE: July 25, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Ian C. McLeod
REGISTRATION NUMBER: 20,931
REFERENCE/DOCKET NUMBER: MSU 4.1-132
TELECOMMUNICATION INFORMATION:
TELEPHONE: (517) 347-4100
TELEFAX: (517) 347-4103
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 91 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: Internal fragment
ORIGINAL SOURCE:
ORGANISM: herpes simplex virus type 1 (HSV1)
NAME/KEY: peptide of HSV1 gene US10 polypeptide
LOCATION: 192 TO 282
OTHER INFORMATION: peptide homologous to the US10 gene
OTHER INFORMATION: polypeptide of MDV
US-09-047-125-10
Query Match 100.0%; Score 21; DB 2; Length 91;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 14
US-07-736-335E-10
Sequence 10, Application US/07736335E
Patent No. 6087127
GENERAL INFORMATION:
APPLICANT: Leland F. Velicer, Peter Brunovskis,
TITLE OF INVENTION: Marek's Disease Herpesvirus
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ian C. McLeod
STREET: 2190 Commons Parkway
CITY: Okemos
STATE: Michigan
COUNTRY: USA
ZIP: 48864
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 1.2 Mb
COMPUTER: IBM PS2, Model 50
OPERATING SYSTEM: MS-DOS 5.0
SOFTWARE: PC-Write 3.02
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/736,335E
FILING DATE: July 25, 1991
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Ian C. McLeod
REGISTRATION NUMBER: 20,931
REFERENCE/DOCKET NUMBER: MSU 4.1-132
TELECOMMUNICATION INFORMATION:
TELEPHONE: (517) 347-4100
TELEFAX: (517) 347-4103
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 91 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal fragment
ORIGINAL SOURCE:
ORGANISM: herpes simplex virus type 1 (HSV1)
FEATURE:
NAME/KEY: peptide of HSV1 gene US10 polypeptide
LOCATION: 192 TO 282
OTHER INFORMATION: peptide homologous to the US10 gene
OTHER INFORMATION: polypeptide of MDV
US-07-736-335E-10

Query Match 100.0%; Score 21; DB 3; Length 91;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HARL 4
DB 30 HARL 33

RESULT 15
US-09-311-352B-2
Sequence 2, Application US/09311352B
Patent No. 6329500
GENERAL INFORMATION:
APPLICANT: Webb, Steven L.
TITLE OF INVENTION: Transforming Growth Factor-beta Binding Site
FILE REFERENCE: 00370-02
CURRENT APPLICATION NUMBER: US/09/311,352B
CURRENT FILING DATE: 1999-05-13
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 113
TYPE: PRT
ORGANISM: Homo sapiens
US-09-311-352B-2

Query Match 100.0%; Score 21; DB 4; Length 113;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HARL 4
DB 19 HARL 22

RESULT 16
US-08-420-235B-29
Sequence 29, Application US/08420235B
Patent No. 5801042
GENERAL INFORMATION:
APPLICANT: Chang, Yuan

APPLICANT: Moore, Patrick S.
TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS
TITLE OF INVENTION: SEQUENCES AND USES THEREOF
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/420,235B
FILING DATE:
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 45185-B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 120 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-420-235B-29

Query Match 100.0%; Score 21; DB 1; Length 120;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HARL 4
DB 100 HARL 103

RESULT 17
US-08-793-624-29
Sequence 29, Application US/08793624C
Patent No. 6150093
GENERAL INFORMATION:
APPLICANT: Chang, Yuan
APPLICANT: Moore, Patrick S.
TITLE OF INVENTION: Unique Associated Kaposi's Sarcoma Virus Sequences And
TITLE OF INVENTION: Uses Thereof
FILE REFERENCE: 45185-C-PCT-US/Jpw
CURRENT APPLICATION NUMBER: US/08/793,624C
CURRENT FILING DATE: 1997-02-18
NUMBER OF SEQ ID NOS: 58
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 29
LENGTH: 120
TYPE: PRT
ORGANISM: Kaposi's sarcoma-associated herpesvirus
US-08-793-624-29

Query Match 100.0%; Score 21; DB 4; Length 120;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HARL 4
DB 100 HARL 103

```
Result 18
PCT-US95-10194-29
; Sequence 29, Application PC/TUS9510194
; GENERAL INFORMATION:
; APPLICANT: The Trustees of Columbia University in the City of New York
; APPLICANT: City
; TITLE OF INVENTION: UNIOQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS
; TITLE OF INVENTION: SEQUENCES AND USES THEREOF
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/10194
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 45185-C-PCT/JPM/MSC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 120 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US95-10194-29

Query Match
Best Local Similarity 100.0%; Score 21; DB 5; Length 120;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HARL 4
DB 100 HARL 103

Result 19
US-09-134-001C-3657
; Sequence 3657, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3657
; LENGTH: 166
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
; US-09-134-001C-3657

Query Match
Best Local Similarity 100.0%; Score 21; DB 4; Length 166;
Matches 100.0%; Pred. No. 2.5e+02;

QY 1 HARL 4
DB 100 HARL 103

Result 20
US-09-186-250-4
; Sequence 4, Application US/09186250A
; Patent No. 6043055
; GENERAL INFORMATION:
; APPLICANT: Hsueh, Aaron
; APPLICANT: Hsu, Sheau
; TITLE OF INVENTION: Bok Genes and their Uses
; FILE REFERENCE: SUN-72P
; CURRENT APPLICATION NUMBER: US/09/186,250A
; CURRENT FILING DATE: 1998-11-04
; EARLIER APPLICATION NUMBER: 60/064,943
; EARLIER FILING DATE: 1997-11-07
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 170
; TYPE: PRT
; ORGANISM: R.rattus
; US-09-186-250-4

Query Match
Best Local Similarity 100.0%; Score 21; DB 3; Length 170;
Matches 100.0%; Pred. No. 2.5e+02;

QY 1 HARL 4
DB 40 HARL 43

Result 21
US-09-186-250-8
; Sequence 8, Application US/09186250A
; Patent No. 6043055
; GENERAL INFORMATION:
; APPLICANT: Hsueh, Aaron
; APPLICANT: Hsu, Sheau
; TITLE OF INVENTION: Bok Genes and their Uses
; FILE REFERENCE: SUN-72P
; CURRENT APPLICATION NUMBER: US/09/186,250A
; CURRENT FILING DATE: 1998-11-04
; EARLIER APPLICATION NUMBER: 60/064,943
; EARLIER FILING DATE: 1997-11-07
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 170
; TYPE: PRT
; ORGANISM: H. sapiens
; US-09-186-250-8

Query Match
Best Local Similarity 100.0%; Score 21; DB 3; Length 170;
Matches 100.0%; Pred. No. 2.5e+02;

QY 1 HARL 4
DB 40 HARL 43

Result 22
US-09-517-347-4
; Sequence 4, Application US/09517347
; Patent No. 6222017
; GENERAL INFORMATION:
; APPLICANT: Hsueh, Aaron
```

```

; APPLICANT: Hsu, Sheau
; TITLE OF INVENTION: Mammalian Pro-Apoptotic Bok Genes and
; FILE OF INVENTION: Their Uses
; FILE REFERENCE: Stan-072DIV
; CURRENT APPLICATION NUMBER: US/09/517,347
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: 09/186,250
; PRIOR FILING DATE: 1998-11-04
; PRIOR APPLICATION NUMBER: 60/064,943
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 170
; TYPE: PRT
; ORGANISM: r.rattus
US-09-517-347-4

Query Match
Best Local Similarity 100.0%; Score 21; DB 4; Length 170;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
    ||||
Db 40 HARL 43

RESULT 23
US-09-517-347-8
; Sequence 8, Application US/09517347
; Patent No. 6222017
; GENERAL INFORMATION:
; APPLICANT: Hsueh, Aaron
; APPLICANT: Hsu, Sheau
; TITLE OF INVENTION: Mammalian Pro-Apoptotic Bok Genes and
; FILE OF INVENTION: Their Uses
; FILE REFERENCE: Stan-072DIV
; CURRENT APPLICATION NUMBER: US/09/517,347
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: 09/186,250
; PRIOR FILING DATE: 1998-11-04
; PRIOR APPLICATION NUMBER: 60/064,943
; PRIOR FILING DATE: 1997-11-07
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 170
; TYPE: PRT
; ORGANISM: H. sapiens
US-09-517-347-8

Query Match
Best Local Similarity 100.0%; Score 21; DB 4; Length 170;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
    ||||
Db 40 HARL 43

RESULT 24
US-09-518-232A-4
; Sequence 4, Application US/09518232A
; Patent No. 6376247
; GENERAL INFORMATION:
; APPLICANT: Hsueh, Aaron J.W.
; APPLICANT: Hsu, Sheau Yu
; TITLE OF INVENTION: Mammalian Pro-Apoptotic Bok Genes and
; FILE OF INVENTION: Their Uses
; FILE REFERENCE: STAN072DIV3
; CURRENT APPLICATION NUMBER: US/09/518,232A
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 09/186,250
; PRIOR FILING DATE: 1997-11-07
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 170
; TYPE: PRT
; ORGANISM: H. sapiens
US-09-518-232A-4

Query Match
Best Local Similarity 100.0%; Score 21; DB 4; Length 170;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
    ||||
Db 40 HARL 43
```

```

; PRIOR FILING DATE: 1998-11-04
; PRIOR APPLICATION NUMBER: 60/064,943
; PRIOR FILING DATE: 1997-11-07
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 170
; TYPE: PRT
; ORGANISM: r.rattus
US-09-518-232A-4

Query Match
Best Local Similarity 100.0%; Score 21; DB 4; Length 170;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
    ||||
Db 40 HARL 43

RESULT 25
US-09-518-232A-8
; Sequence 8, Application US/09518232A
; Patent No. 6376247
; GENERAL INFORMATION:
; APPLICANT: Hsueh, Aaron J.W.
; APPLICANT: Hsu, Sheau Yu
; TITLE OF INVENTION: Mammalian Pro-Apoptotic Bok Genes and
; FILE OF INVENTION: Their Uses
; FILE REFERENCE: STAN072DIV3
; CURRENT APPLICATION NUMBER: US/09/518,232A
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 09/186,250
; PRIOR FILING DATE: 1998-11-04
; PRIOR APPLICATION NUMBER: 60/064,943
; PRIOR FILING DATE: 1997-11-07
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 170
; TYPE: PRT
; ORGANISM: H. sapiens
US-09-518-232A-8

Query Match
Best Local Similarity 100.0%; Score 21; DB 4; Length 170;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
    ||||
Db 40 HARL 43

RESULT 26
US-09-517-358A-4
; Sequence 4, Application US/09517358A
; Patent No. 6437097
; GENERAL INFORMATION:
; APPLICANT: Hsueh, Aaron J.W.
; APPLICANT: Hsu, Sheau Yu
; TITLE OF INVENTION: Mammalian Pro-Apoptotic Bok Genes and
; FILE OF INVENTION: Their Uses
; FILE REFERENCE: STAN-072DIV2
; CURRENT APPLICATION NUMBER: US/09/517,358A
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: 09/186,250
; PRIOR FILING DATE: 1998-11-04
; PRIOR APPLICATION NUMBER: 60/064,943
; PRIOR FILING DATE: 1997-11-07
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 170
; TYPE: PRT
; ORGANISM: H. sapiens
US-09-517-358A-4

Query Match
Best Local Similarity 100.0%; Score 21; DB 4; Length 170;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
    ||||
Db 40 HARL 43
```

TYPE: PRT
ORGANISM: R. rattus
US-09-517-358A-4

Query Match 100.0%; Score 21; DB 4; Length 170;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HARL 4
|||||
DB 40 HARL 43

RESULT 27
US-09-517-358A-8
Sequence 8, Application US/09517358A
Patent No. 6437097
GENERAL INFORMATION:
APPLICANT: Hsueh, Aaron J.W.
APPLICANT: Hsu, Sheau Yu
TITLE OF INVENTION: Mammalian Pro-Apoptotic Bcl Genes and
FILE REFERENCE: STAN-072DIV2
CURRENT APPLICATION NUMBER: US/09/517,358A
CURRENT FILING DATE: 2000-03-02
PRIOR APPLICATION NUMBER: 09/186,250
PRIOR FILING DATE: 1998-11-04
PRIOR APPLICATION NUMBER: 60/064,943
PRIOR FILING DATE: 1997-11-07
NUMBER OF SEQ ID NOS: 18
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 8
LENGTH: 170
TYPE: PRT
ORGANISM: H. sapiens
US-09-517-358A-8

Query Match 100.0%; Score 21; DB 4; Length 170;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HARL 4
|||||
DB 40 HARL 43

RESULT 28
US-08-469-537A-84
Sequence 84, Application US/08469537A
Patent No. 5843749
GENERAL INFORMATION:
APPLICANT: Maisongierre, et al.
TITLE OF INVENTION: EHK AND ROR TYROSINE
NUMBER OF SEQUENCES: 107
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Regeneron Pharmaceuticals, Inc.
STREET: 777 Old Saw Mill River Road
CITY: Tarrytown
STATE: NY
COUNTRY: U.S.A.
ZIP: 10591
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469,537A
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 08/406,247

FILING DATE: 17-MAR-1995
APPLICATION NUMBER: USSN 08/144,992
FILING DATE: 28-OCT-1993
APPLICATION NUMBER: USSN 07/736,559
FILING DATE: 26-JUL-1991
ATTORNEY/AGENT INFORMATION:
NAME: Kempster, Ph.D., Gail M
REGISTRATION NUMBER: 32,143
REFERENCE/DOCKET NUMBER: REG 070C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 914-345-7400
TELEFAX: 914-345-7721
TELEX:
INFORMATION FOR SEQ ID NO: 84:
SEQUENCE CHARACTERISTICS:
LENGTH: 179 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-469-537A-84

Query Match 100.0%; Score 21; DB 2; Length 179;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HARL 4
|||||
DB 174 HARL 177

RESULT 29
US-09-311-352B-4
Sequence 4, Application US/09311352B
Patent No. 6329500
GENERAL INFORMATION:
APPLICANT: Webb, Donna J.
APPLICANT: Gonias, Steven L.
TITLE OF INVENTION: Transforming Growth Factor-beta Binding Site
FILE REFERENCE: 00370-02
CURRENT APPLICATION NUMBER: US/09/311,352B
CURRENT FILING DATE: 1999-05-13
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 183
TYPE: PRT
ORGANISM: Homo sapiens
US-09-311-352B-4

Query Match 100.0%; Score 21; DB 4; Length 183;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HARL 4
|||||
DB 104 HARL 107

RESULT 30
US-08-469-537A-37
Sequence 37, Application US/08469537A
Patent No. 5843749
GENERAL INFORMATION:
APPLICANT: Maisongierre, et al.
TITLE OF INVENTION: EHK AND ROR TYROSINE
NUMBER OF SEQUENCES: 107
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Regeneron Pharmaceuticals, Inc.
STREET: 777 Old Saw Mill River Road
CITY: Tarrytown
STATE: NY

COUNTRY: U.S.A.
ZIP: 10591
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469,537A
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 08/406,247
FILING DATE: 17-MAR-1995
APPLICATION NUMBER: USSN 08/144,992
FILING DATE: 28-OCT-1993
APPLICATION NUMBER: USSN 07/736,559
FILING DATE: 26-JUL-1991
ATTORNEY/AGENT INFORMATION:
NAME: Kempster, P.D., Gail M
REGISTRATION NUMBER: 32,143
REFERENCE/DOCKET NUMBER: REG 070C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 914-345-7400
TELEFAX: 914-345-7721
TELEX:
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 192 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-469-537A-37

Query Match 100.0%; Score 21; DB 2; Length 192;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
|||
Db 174 HARL 177

RESULT 31
US-07-661-610C-5
Sequence 5, Application US/07661610C
Patent No. 5292643
GENERAL INFORMATION:
APPLICANT: Shidano, Yuji
APPLICANT: Toyoda, Hideyoshi
APPLICANT: Utsuni, Ryutaro
APPLICANT: Obata, Kazuaki
TITLE OF INVENTION: Fusaric Acid Resistant Genes
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MATER & NEUSTADT,
ADDRESSER: P.C.
STREET: 1755 Jefferson Davis Highway, Fourth Floor
CITY: Arlington
STATE: Virginia
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/661,610C
FILING DATE: 19910228
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Oblon, No. 5292643man F.

REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 2292-010-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)521-4500
TELEFAX: (703)486-2347
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 208 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-661-610C-5

Query Match 100.0%; Score 21; DB 1; Length 208;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
|||
Db 46 HARL 49

RESULT 32
US-08-309-512-4
Sequence 4, Application US/08309512
Patent No. 5759828
GENERAL INFORMATION:
APPLICANT: Tal, Ronny
APPLICANT: Benziman, Moshe
APPLICANT: Gelfand, David H.
APPLICANT: Ben-Bassat, Arle
APPLICANT: Calhoun, Roger D.
APPLICANT: Mong, Hing C.
TITLE OF INVENTION: CYCLIC DIGUANYLATE METABOLIC ENZYMES
NUMBER OF SEQUENCES: 63
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 2730 Sand Hill Road
CITY: Menlo Park
STATE: California
COUNTRY: U.S.A.
ZIP: 94025
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/309,512
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/800,218
FILING DATE: 29-NOV-1991
ATTORNEY/AGENT INFORMATION:
NAME: Bortner, Scott R.
REGISTRATION NUMBER: 34,298
REFERENCE/DOCKET NUMBER: 8145-008
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 854-3660
TELEFAX: (415) 854-3694
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 208 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Acetobacter xylinum

US-08-309-512-4

Query Match

Best Local Similarity 100.0%; Score 21; DB 1; Length 208;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;QY 1 HARL 4
DB 118 HARL 121

RESULT 33

PCT-US92-08756A-4
Sequence 4, Application PC/TUS9208756A

GENERAL INFORMATION:

APPLICANT: Tal, Ronny
APPLICANT: Benzman, Moshe
APPLICANT: Gelfand, David H.
APPLICANT: Ben-Bassat, Arie
APPLICANT: Calhoun, Roger D.
APPLICANT: Wong, Hing C.
TITLE OF INVENTION: CYCLIC DIGUANYLATE
TITLE OF INVENTION: METABOLIC ENZYMES
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Limbach and Limbach
STREET: 2001 Ferry Building
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0,
SOFTWARE: Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/08756A
FILING DATE: 19921014
CLASSIFICATION: 435

Prior Application DATA:

APPLICATION NUMBER: US 07/800,218
FILING DATE: 29-NOV-1991

ATTORNEY/AGENT INFORMATION:

NAME: Portner, Scott R.

REGISTRATION NUMBER: 34,298

REFERENCE/DOCKET NUMBER: WEYR 20050 USA

TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-433-4150

TELEFAX: 415-433-8716

TELEX: 278356

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:
LENGTH: 208 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: NO
AMTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Acetobacter xylinum
PCT-US92-08756A-4

Query Match

Best Local Similarity 100.0%; Score 21; DB 5; Length 208;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;QY 1 HARL 4
DB 118 HARL 121

RESULT 34

US-09-199-637A-401
Sequence 401, Application US/09199637A

Patent No. 6355411

GENERAL INFORMATION:

APPLICANT: Ausubel, Frederick
APPLICANT: Goodman, Howard M.
APPLICANT: Rahme, Laurence G.
APPLICANT: Mahajan-Miklos, Shalina
APPLICANT: Tan, Man-Wah
APPLICANT: Cao, Hui
APPLICANT: Drenkard, Eliana
APPLICANT: Tsongalis, John
TITLE OF INVENTION: VIRULENCE-ASSOCIATED NUCLEIC ACID
TITLE OF INVENTION: SEQUENCES AND USES THEREOF
FILE REFERENCE: 00786/361002
CURRENT APPLICATION NUMBER: US/09/199,637A
CURRENT FILING DATE: 1998-11-25
PRIOR APPLICATION NUMBER: 60/066,517
PRIOR FILING DATE: 1997-11-25
NUMBER OF SEQ ID NOS: 437
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 401
LENGTH: 209
TYPE: PRTORGANISM: Pseudomonas aeruginosa
US-09-199-637A-401

Query Match

Best Local Similarity 100.0%; Score 21; DB 4; Length 209;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;QY 1 HARL 4
DB 203 HARL 206

RESULT 35

US-09-186-250-2
Sequence 2, Application US/09186250A

Patent No. 6043055

GENERAL INFORMATION:

APPLICANT: Hsueh, Aaron
APPLICANT: Hsu, Sheau
TITLE OF INVENTION: BOK Genes and their Uses
FILE REFERENCE: SUN-72P
CURRENT APPLICATION NUMBER: US/09/186,250A
CURRENT FILING DATE: 1998-11-04
EARLIER APPLICATION NUMBER: 60/064,943
EARLIER FILING DATE: 1997-11-07
NUMBER OF SEQ ID NOS: 18
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 213
TYPE: PRT
ORGANISM: R. rattus
US-09-186-250-2

Query Match

Best Local Similarity 100.0%; Score 21; DB 3; Length 213;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;QY 1 HARL 4
DB 40 HARL 43

RESULT 36

US-09-186-250-6
Sequence 6, Application US/09186250A

GENERAL INFORMATION:

APPLICANT: Hsueh, Aaron


```

; APPLICANT: Hsu, Sheau
; TITLE OF INVENTION: Bok Genes and their Uses
; FILE REFERENCE: SUN-72P
; CURRENT APPLICATION NUMBER: US/09/186,250A
; CURRENT FILING DATE: 1998-11-04
; EARLIER APPLICATION NUMBER: 60/064,943
; EARLIER FILING DATE: 1997-11-07
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 213
; TYPE: PRT
; ORGANISM: H.sapiens
US-09-186-250-6

Query Match
Best Local Similarity 100.0%; Score 21; DB 3; Length 213;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HARL 4
    ||||
Db 40 HARL 43

RESULT 37
US-09-517-347-2
; Sequence 2, Application US/09517347
; Patent No. 6222017
; GENERAL INFORMATION:
; APPLICANT: Hsueh, Aaron
; TITLE OF INVENTION: Mammalian Pro-Apoptotic Bok Genes and
; FILE REFERENCE: Stan-072DIV
; CURRENT APPLICATION NUMBER: US/09/517,347
; CURRENT FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: 09/186,250
; PRIOR FILING DATE: 1998-11-04
; PRIOR APPLICATION NUMBER: 60/064,943
; PRIOR FILING DATE: 1997-11-07
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 213
; TYPE: PRT
; ORGANISM: R. rattus
US-09-517-347-2

Query Match
Best Local Similarity 100.0%; Score 21; DB 4; Length 213;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HARL 4
    ||||
Db 40 HARL 43

RESULT 38
US-09-517-347-6
; Sequence 6, Application US/09517347
; Patent No. 6222017
; GENERAL INFORMATION:
; APPLICANT: Hsueh, Aaron
; TITLE OF INVENTION: Mammalian Pro-Apoptotic Bok Genes and
; FILE REFERENCE: Stan-072DIV
; CURRENT APPLICATION NUMBER: US/09/517,347
; CURRENT FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: 09/186,250
; PRIOR FILING DATE: 1998-11-04
; PRIOR APPLICATION NUMBER: 60/064,943
; PRIOR FILING DATE: 1997-11-07
```

```

; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 213
; TYPE: PRT
; ORGANISM: H.sapiens
US-09-517-347-6

Query Match
Best Local Similarity 100.0%; Score 21; DB 4; Length 213;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HARL 4
    ||||
Db 40 HARL 43

RESULT 39
US-09-518-232A-2
; Sequence 2, Application US/09518232A
; Patent No. 6376247
; GENERAL INFORMATION:
; APPLICANT: Hsueh, Aaron J.W.
; TITLE OF INVENTION: Mammalian Pro-Apoptotic Bok Genes and
; FILE REFERENCE: STAN072DIV3
; CURRENT APPLICATION NUMBER: US/09/518,232A
; CURRENT FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 09/186,250
; PRIOR FILING DATE: 1998-11-04
; PRIOR APPLICATION NUMBER: 60/064,943
; PRIOR FILING DATE: 1997-11-07
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 213
; TYPE: PRT
; ORGANISM: R. rattus
US-09-518-232A-2

Query Match
Best Local Similarity 100.0%; Score 21; DB 4; Length 213;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HARL 4
    ||||
Db 40 HARL 43

RESULT 40
US-09-518-232A-6
; Sequence 6, Application US/09518232A
; Patent No. 6376247
; GENERAL INFORMATION:
; APPLICANT: Hsueh, Aaron J.W.
; TITLE OF INVENTION: Mammalian Pro-Apoptotic Bok Genes and
; FILE REFERENCE: STAN072DIV3
; CURRENT APPLICATION NUMBER: US/09/518,232A
; CURRENT FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 09/186,250
; PRIOR FILING DATE: 1998-11-04
; PRIOR APPLICATION NUMBER: 60/064,943
; PRIOR FILING DATE: 1997-11-07
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 213
; TYPE: PRT
; ORGANISM: H.sapiens
US-09-518-232A-6
```

Query Match 100.0%; Score 21; DB 4; Length 213;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HARL 4
||||
DB 40 HARL 43

RESULT 41
US-09-517-358A-2
; Sequence 2, Application US/09517358A
; Patent No. 6437097
; GENERAL INFORMATION:
; APPLICANT: Hsueh, Aaron J.W.
; TITLE OF INVENTION: Mammalian Pro-Apoptotic Bcl Genes and
; FILE REFERENCE: STAN-072DIV2
; CURRENT APPLICATION NUMBER: US/09/517,358A
; CURRENT FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: 09/186,250
; PRIOR FILING DATE: 1998-11-04
; PRIOR APPLICATION NUMBER: 60/064,943
; PRIOR FILING DATE: 1997-11-07
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 213
; TYPE: PRT
; ORGANISM: r. rattus
US-09-517-358A-2

Query Match 100.0%; Score 21; DB 4; Length 213;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HARL 4
||||
DB 40 HARL 43

RESULT 42
US-09-517-358A-6
; Sequence 6, Application US/09517358A
; Patent No. 6437097
; GENERAL INFORMATION:
; APPLICANT: Hsueh, Aaron J.W.
; APPLICANT: Hsu, Sheau Yu
; TITLE OF INVENTION: Mammalian Pro-Apoptotic Bcl Genes and
; FILE REFERENCE: STAN-072DIV2
; CURRENT APPLICATION NUMBER: US/09/517,358A
; CURRENT FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: 09/186,250
; PRIOR FILING DATE: 1998-11-04
; PRIOR APPLICATION NUMBER: 60/064,943
; PRIOR FILING DATE: 1997-11-07
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 213
; TYPE: PRT
; ORGANISM: H.sapiens
US-09-517-358A-6

Query Match 100.0%; Score 21; DB 4; Length 213;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HARL 4
||||

DB 40 HARL 43

RESULT 43
US-08-796-676-3
; Sequence 3, Application US/08796676
; Patent No. 5858712
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Goli, Surya K.
; TITLE OF INVENTION: HUMAN LEA-MOTIF DEVELOPMENTAL
; TITLE OF INVENTION: . PROTEIN
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/796,676
; FILING DATE: Filed Herewith
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0213 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 215 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 969170
US-08-796-676-3

Query Match 100.0%; Score 21; DB 2; Length 215;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HARL 4
||||
DB 102 HARL 105

RESULT 44
US-08-967-364-8
; Sequence 8, Application US/08967364
; Patent No. 5989859
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Lal, Preeti
; APPLICANT: Guegler, Karl J.
; APPLICANT: Shah, Purvi
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: VESICLE TRAFFICKING PROTEINS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA

COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/967,364
FILING DATE: NO. 598959ember 7, 1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Cerione, Michael C.
REGISTRATION NUMBER: 39,132
REFERENCE/DOCKET NUMBER: PF-0417 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 215 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: G1969170
US-08-967-364-8

Query Match 100.0%; Score 21; DB 2; Length 215;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HARL 4
DB 102 HARL 105

RESULT 45
US-09-368-408-8
Sequence 8, Application US/09368408
Patent No. 6071703
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Lal, Preeti
APPLICANT: Guegler, Karl J.
APPLICANT: Shah, Purvi
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: VESICLE TRAFFICKING PROTEINS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Dr.
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/368,408
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/967,364
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Cerione, Michael C.
REGISTRATION NUMBER: 39,132

REFERENCE/DOCKET NUMBER: PF-0417 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 215 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: G1969170
US-09-368-408-8

Query Match 100.0%; Score 21; DB 3; Length 215;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HARL 4
DB 102 HARL 105

RESULT 46
US-09-213-391-3
Sequence 3, Application US/09213391
Patent No. 6281190
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Goll, Surya K.
TITLE OF INVENTION: HUMAN LEA-MOTIF DEVELOPMENTAL
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/213,391
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/796,676
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0213 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 215 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: G669170
US-09-213-391-3

Query Match 100.0%; Score 21; DB 4; Length 215;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
||||

Db 102 HARL 105

RESULT 47

US-08-796-676-1

; Sequence 1, Application US/08796676

; Patent No. 5858712

; GENERAL INFORMATION:

; APPLICANT: Hillman, Jennifer L.

; TITLE OF INVENTION: HUMAN LEA-MOTIF DEVELOPMENTAL

; TITLE OF INVENTION: PROTEIN

; NUMBER OF SEQUENCES: 3

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Incyte Pharmaceuticals, Inc.

; STREET: 3174 Porter Drive

; CITY: Palo Alto

; STATE: CA

; COUNTRY: USA

; ZIP: 94304

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/796,676

; FILING DATE: Filed Herewith

; CLASSIFICATION: 424

; ATTORNEY/AGENT INFORMATION:

; NAME: Billings, Lucy J.

; REGISTRATION NUMBER: 36,749

; REFERENCE/DOCKET NUMBER: PF-0213 US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-855-0555

; TELEFAX: 415-845-4166

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 219 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; US-08-796-676-1

Query Match 100.0%; Score 21; DB 2; Length 219;

Best Local Similarity 100.0%; Pred. No. 3.3e+02; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
||||

Db 102 HARL 105

RESULT 48

US-09-213-391-1

; Sequence 1, Application US/09213391

; Patent No. 6281190

; GENERAL INFORMATION:

; APPLICANT: Hillman, Jennifer L.

; TITLE OF INVENTION: HUMAN LEA-MOTIF DEVELOPMENTAL

; TITLE OF INVENTION: PROTEIN

; NUMBER OF SEQUENCES: 3

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Incyte Pharmaceuticals, Inc.

; STREET: 3174 Porter Drive

; CITY: Palo Alto

; STATE: CA

; COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/213,391

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/796,676

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0213 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-855-0555

TELEFAX: 415-845-4166

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 219 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

US-09-213-391-1

Query Match 100.0%; Score 21; DB 4; Length 219;

Best Local Similarity 100.0%; Pred. No. 3.3e+02; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
||||

Db 102 HARL 105

RESULT 49

US-09-166-350-20

; Sequence 20, Application US/09166350A

; Patent No. 6440663

; GENERAL INFORMATION:

; APPLICANT: Scanlan, Matthew

; APPLICANT: Chen, Yao

; APPLICANT: Stockert, Elisabeth

; APPLICANT: Old, Lloyd

; APPLICANT: Jager, Elke

; TITLE OF INVENTION: Renal Cancer Associated Antigens and

; FILE REFERENCE: L0461/7051

; CURRENT APPLICATION NUMBER: US/09/166,350A

; CURRENT FILING DATE: 1998-10-05

; EARLIER APPLICATION NUMBER: US 09/166,350

; EARLIER FILING DATE: 1998-10-05

; NUMBER OF SEQ ID NOS: 35

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 20

; LENGTH: 223

; TYPE: PRT

; ORGANISM: Homo sapiens

; US-09-166-350-20

Query Match 100.0%; Score 21; DB 4; Length 223;

Best Local Similarity 100.0%; Pred. No. 3.3e+02; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
||||

Db 191 HARL 194

RESULT 50

```

US-09-161-241-14
; Sequence 14, Application US/09161241
; Patent No. 6344541
; GENERAL INFORMATION:
; APPLICANT: Bass, Michael B
; APPLICANT: Sullivan, John K
; APPLICANT: Theill, Lars E
; APPLICANT: Wang, Daquang
; TITLE OF INVENTION: NOVEL DKR POLYPEPTIDES
; FILE REFERENCE: A-548
; CURRENT APPLICATION NUMBER: US/09/161,241
; CURRENT FILING DATE: 1998-09-25
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: Patentn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 224
; TYPE: PRT
; ORGANISM: Human
US-09-161-241-14

```

```

Query Match      100.0%; Score 21; DB 4; Length 224;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 HARL 4
        ||||
DB      212 HARL 215

```

Search completed: November 13, 2002, 13:42:36
 Job time : 15 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2002, 13:41:44 ; Search time 10 Seconds
(without alignments)
6.024 Million cell updates/sec

Title: US-09-697-590-2_COPY_292_295

Perfect score: 21

Sequence: 1 HARL 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 97044 seqs, 15060890 residues

Total number of hits satisfying chosen parameters: 97044

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA:*

1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCF_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/PCFOS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	21	100.0	32	10	US-09-864-761-37322
2	21	100.0	53	10	US-09-982-610-6
3	21	100.0	68	10	US-09-864-761-33341
4	21	100.0	69	10	US-09-864-761-46505
5	21	100.0	76	10	US-09-925-300-1130
6	21	100.0	94	10	US-09-729-674-162
7	21	100.0	117	10	US-09-925-299-1124
8	21	100.0	135	10	US-09-800-729-194
9	21	100.0	155	9	US-09-712-363-155
10	21	100.0	170	10	US-09-682-667-4
11	21	100.0	170	10	US-09-682-667-8
12	21	100.0	183	10	US-09-731-872-357
13	21	100.0	186	10	US-09-815-242-11908
14	21	100.0	213	10	US-09-682-667-2
15	21	100.0	213	10	US-09-682-667-6
16	21	100.0	216	10	US-09-750-373-22
17	21	100.0	224	9	US-09-976-736-14
18	21	100.0	293	10	US-09-323-998D-15
19	21	100.0	293	10	US-09-323-998D-43

20	21	100.0	298	10	US-09-815-242-13325	Sequence 13325, A
21	21	100.0	298	10	US-09-815-242-13603	Sequence 13603, A
22	21	100.0	305	10	US-09-323-998D-14	Sequence 14, Appl
23	21	100.0	305	10	US-09-323-998D-44	Sequence 44, Appl
24	21	100.0	312	10	US-09-815-242-4904	Sequence 4904, Ap
25	21	100.0	312	10	US-09-815-242-10541	Sequence 10541, A
26	21	100.0	331	10	US-09-815-242-11761	Sequence 11761, A
27	21	100.0	350	9	US-09-361-655-14	Sequence 14, Appl
28	21	100.0	352	10	US-09-860-351-2	Sequence 2, Appl
29	21	100.0	358	10	US-09-815-242-12071	Sequence 12071, A
30	21	100.0	359	10	US-09-815-242-51116	Sequence 5116, Ap
31	21	100.0	370	10	US-09-833-790-253	Sequence 253, App
32	21	100.0	372	9	US-09-973-941-4	Sequence 4, Appl
33	21	100.0	372	10	US-09-973-963-4	Sequence 4, Appl
34	21	100.0	372	10	US-09-973-064-4	Sequence 4, Appl
35	21	100.0	372	10	US-09-973-077-4	Sequence 4, Appl
36	21	100.0	372	10	US-09-973-063-4	Sequence 4, Appl
37	21	100.0	372	10	US-09-973-964-4	Sequence 4, Appl
38	21	100.0	372	10	US-09-975-072-4	Sequence 4, Appl
39	21	100.0	372	10	US-09-972-038-4	Sequence 4, Appl
40	21	100.0	372	10	US-09-972-757-4	Sequence 4, Appl
41	21	100.0	372	10	US-09-973-965-4	Sequence 4, Appl
42	21	100.0	373	10	US-09-925-300-1655	Sequence 1655, Ap
43	21	100.0	375	10	US-09-964-666-2	Sequence 2, Appl
44	21	100.0	375	10	US-09-964-412-2	Sequence 2, Appl
45	21	100.0	375	10	US-09-371-900-38	Sequence 38, Appl

ALIGNMENTS

RESULT 1
US-09-864-761-37322
Sequence 37322, Application US/09864761
Patent No. US20020046763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharron G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL I
FILE REFERENCE: Aeonlica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 37322
LENGTH: 32
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AL049633.3
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.2
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.3
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.6
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 3.5
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.1
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4.1
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 5
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.9
OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 3.5
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 3.3
US-09-864-761-37322

Query Match 100.0%; Score 21; DB 10; Length 32;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
||||
DB 12 HARL 15

RESULT 2

US-09-982-610-6
Sequence 6, Application US/09982610
Patent No. US20020146420A1

GENERAL INFORMATION:

APPLICANT: Genentech, Inc.
Benneet, Brian D.
Goeddel, David
Lee, James M.
Mathews, William
Tsai, Siao Ping
Wood, William I.
TITLE OF INVENTION: PROTEIN TYROSINE KINASE AGONIST ANTIBODIES
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 1.44 mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Winpatin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/982,610
FILING DATE: 17-Oct-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/446,648
FILING DATE: 1996-MAY-23
APPLICATION NUMBER: 08/222616
FILING DATE: 04-APR-1994

ATTORNEY/AGENT INFORMATION:

NAME: Lee, Wendy M.
REGISTRATION NUMBER: 40,378
REFERENCE/DOCKET NUMBER: P0821P3PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 53 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-982-610-6

Query Match 100.0%; Score 21; DB 10; Length 53;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
||||
DB 41 HARL 44

RESULT 3

US-09-864-761-33341
Sequence 33341, Application US/09864761
Patent No. US20020048763A1

GENERAL INFORMATION:

APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.

APPLICANT: Chen, Wensheng

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL F

FILE REFERENCE: Aecomica-X-1

CURRENT APPLICATION NUMBER: US/09/864,761

CURRENT FILING DATE: 2001-05-23

PRIOR APPLICATION NUMBER: US 60/180,312

PRIOR FILING DATE: 2000-02-04

PRIOR APPLICATION NUMBER: US 60/207,456

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: US 09/632,366

PRIOR FILING DATE: 2000-08-03

PRIOR APPLICATION NUMBER: GB 24263.6

PRIOR FILING DATE: 2000-10-04

PRIOR APPLICATION NUMBER: US 60/236,359

PRIOR FILING DATE: 2000-09-27

PRIOR APPLICATION NUMBER: PCT/US01/00666

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00667

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00664

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00669

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00665

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00668

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00663

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00662

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00661

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00670

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: US 60/234,687

PRIOR FILING DATE: 2000-09-21

PRIOR APPLICATION NUMBER: US 09/608,408

PRIOR FILING DATE: 2000-06-30

PRIOR APPLICATION NUMBER: US 09/774,203

```
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 33341
; LENGTH: 68
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC000353.21
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN HB1100, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.8
; OTHER INFORMATION: EST_HUMAN HIT: A1934403.1, EVALU 3.00e-07
; OTHER INFORMATION: SWISSPROT HIT: O76082, EVALU 1.00e-06
US-09-864-761-33341

Query Match          100.0%; Score 21; DB 10; Length 68;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
    ||||
Db 10 HARL 13

RESULT 4
US-09-864-761-46505
; Sequence 46505, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecomica-X-1
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
```

```
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 46505
; LENGTH: 69
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL034343.17
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 3.6
; OTHER INFORMATION: EXPRESSED IN HB1100, SIGNAL = 5.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 4.9
; OTHER INFORMATION: SWISSPROT HIT: P23206, EVALU 4.00e-34
; OTHER INFORMATION: EST_HUMAN HIT: A0124791.1, EVALU 5.00e-33
US-09-864-761-46505

Query Match          100.0%; Score 21; DB 10; Length 69;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
    ||||
Db 50 HARL 53

RESULT 5
US-09-925-300-1130
; Sequence 1130, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: US/09/925,300
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1130
; LENGTH: 76
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-300-1130

Query Match          100.0%; Score 21; DB 10; Length 76;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
    ||||
Db 25 HARL 28

RESULT 6
US-09-729-674-162
; Sequence 162, Application US/09729674
; Patent No. US20010039335A1
```


GENERAL INFORMATION:
APPLICANT: Jacobs, Kenneth
APPLICANT: McCoy, John M.
APPLICANT: Lavallee, Edward R.
APPLICANT: Collins-Racie, Lisa A.
APPLICANT: Evans, Cheryl
APPLICANT: Merberg, David
APPLICANT: Treacy, Maurice
APPLICANT: Agostino, Michael J.
APPLICANT: Steininger II, Robert J.
APPLICANT: Spaulding, Vikki
APPLICANT: Wong, Gordon G.
APPLICANT: Clark, Hilary
APPLICANT: Fechtel, Kim
APPLICANT: Genetics Institute, Inc.
TITLE OF INVENTION: SECRETED PROTEINS AND POLYPEPTIDES ENCODING THEM
FILE REFERENCE: 6055-64X
CURRENT APPLICATION NUMBER: US/09/729,674
CURRENT FILING DATE: 2000-12-04
PRIOR APPLICATION NUMBER: 09/539,330
PRIOR FILING DATE: 2000-03-30
NUMBER OF SEQ ID NOS: 283
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 162
LENGTH: 94
TYPE: PRT
ORGANISM: Homo sapiens
US-09-729-674-162

Query Match 100.0%; Score 21; DB 10; Length 94;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HARL 4
1111
DB 30 HARL 33

RESULT 7
US-09-925-299-1124
Sequence 1124, Application US/09925299
Patent No. US20020055627A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA102
CURRENT APPLICATION NUMBER: US/09/925,299
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05883
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 1556
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1124
LENGTH: 117
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (7)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (87)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (97)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (99)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (110)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-299-1124

Query Match 100.0%; Score 21; DB 10; Length 117;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HARL 4
1111
DB 81 HARL 84

RESULT 8
US-09-800-729-194
Sequence 194, Application US/09800729
Patent No. US20020068319A1
GENERAL INFORMATION:
APPLICANT: Ni et al.
TITLE OF INVENTION: 32 Human secreted proteins
FILE REFERENCE: P2044P1
CURRENT APPLICATION NUMBER: US/09/800,729
CURRENT FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: PCT/US00/26013
PRIOR FILING DATE: 2000-09-22
PRIOR APPLICATION NUMBER: 60/155,709
PRIOR FILING DATE: 1999-09-24
NUMBER OF SEQ ID NOS: 217
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 194
LENGTH: 135
TYPE: PRT
ORGANISM: Homo sapiens
US-09-800-729-194

Query Match 100.0%; Score 21; DB 10; Length 135;
Best Local Similarity 100.0%; Pred. No. 98;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HARL 4
1111
DB 60 HARL 63

RESULT 9
US-09-712-363-155
Sequence 155, Application US/09712363
Patent No. US20020164588A1
GENERAL INFORMATION:
APPLICANT: Eisenberg, David
APPLICANT: Rotstein, Sergio H.
APPLICANT: Marcotte, Edward M.
TITLE OF INVENTION: DETERMINING THE FUNCTIONS AND
FILE REFERENCE: 07419-032001
CURRENT APPLICATION NUMBER: US/09/712,363
CURRENT FILING DATE: 2000-11-13
PRIOR APPLICATION NUMBER: PCT/US00/02246
PRIOR FILING DATE: 2000-01-28
PRIOR APPLICATION NUMBER: 60/179,531
PRIOR FILING DATE: 2000-02-01
PRIOR APPLICATION NUMBER: 60/117,844
PRIOR FILING DATE: 1999-01-29
PRIOR APPLICATION NUMBER: 60/118,206
PRIOR FILING DATE: 1999-02-01
PRIOR APPLICATION NUMBER: 60/126,593
PRIOR FILING DATE: 1999-03-26
PRIOR APPLICATION NUMBER: 60/134,093
PRIOR FILING DATE: 1999-05-14
PRIOR APPLICATION NUMBER: 60/134,092
PRIOR FILING DATE: 1999-05-14
PRIOR APPLICATION NUMBER: 60/165,124
PRIOR FILING DATE: 1999-11-12
PRIOR APPLICATION NUMBER: 60/165,086

;; PRIOR FILING DATE: 1999-11-12
;; NUMBER OF SEQ ID NOS: 292
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 155
;; LENGTH: 155
;; TYPE: PRT
;; ORGANISM: Mycobacterium tuberculosis
US-09-712-363-155

Query Match
Best Local Similarity 100.0%; Score 21; DB 9; Length 155;
Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
Db 104 HARL 107

RESULT 10
US-09-682-667-4
;; Sequence 4, Application US/09682667
;; Patent No. US20020076794A1
;; GENERAL INFORMATION:
;; APPLICANT: Hsueh, Aaron J.W.
;; TITLE OF INVENTION: MAMMALIAN PRO-APOPTOTIC BOK GENES AND
;; FILE REFERENCE: STAN072CON
;; CURRENT APPLICATION NUMBER: US/09/682,667
;; PRIOR FILING DATE: 2001-10-05
;; PRIOR APPLICATION NUMBER: 09/517,358
;; PRIOR FILING DATE: 2000-03-02
;; PRIOR APPLICATION NUMBER: 09/186,250
;; PRIOR FILING DATE: 1998-11-04
;; PRIOR APPLICATION NUMBER: 60/064,943
;; PRIOR FILING DATE: 1997-11-07
;; NUMBER OF SEQ ID NOS: 18
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 4
;; LENGTH: 170
;; TYPE: PRT
;; ORGANISM: R.rattus
US-09-682-667-4

Query Match
Best Local Similarity 100.0%; Score 21; DB 10; Length 170;
Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
Db 40 HARL 43

RESULT 11
US-09-682-667-8
;; Sequence 8, Application US/09682667
;; Patent No. US20020076794A1
;; GENERAL INFORMATION:
;; APPLICANT: Hsueh, Aaron J.W.
;; APPLICANT: Hsu, Sheau Yu
;; TITLE OF INVENTION: MAMMALIAN PRO-APOPTOTIC BOK GENES AND
;; FILE REFERENCE: STAN072CON
;; CURRENT APPLICATION NUMBER: US/09/682,667
;; PRIOR FILING DATE: 2001-10-05
;; PRIOR APPLICATION NUMBER: 09/517,358
;; PRIOR FILING DATE: 2000-03-02
;; PRIOR APPLICATION NUMBER: 09/186,250
;; PRIOR FILING DATE: 1998-11-04
;; PRIOR APPLICATION NUMBER: 60/064,943
;; PRIOR FILING DATE: 1997-11-07
;; NUMBER OF SEQ ID NOS: 18
;; SOFTWARE: FastSeq for Windows Version 4.0

;; SEQ ID NO 8
;; LENGTH: 170
;; TYPE: PRT
;; ORGANISM: H. sapiens
US-09-682-667-8

Query Match
Best Local Similarity 100.0%; Score 21; DB 10; Length 170;
Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
Db 40 HARL 43

RESULT 12
US-09-731-872-357
;; Sequence 357, Application US/09731872
;; Patent No. US20020102604A1
;; GENERAL INFORMATION:
;; APPLICANT: Dumas Milne Edwards, Jean Baptiste
;; APPLICANT: Bougueleret, Lydie
;; APPLICANT: Jobert, Severin
;; TITLE OF INVENTION: FULL-LENGTH HUMAN CDNAS ENCODING POTENTIALLY SECRETED PROTEIN
;; FILE REFERENCE: 78.US3.REG
;; CURRENT APPLICATION NUMBER: US/09/731,872
;; PRIOR FILING DATE: 2000-12-07
;; PRIOR APPLICATION NUMBER: US 60/169,629
;; PRIOR FILING DATE: 1999-12-08
;; PRIOR APPLICATION NUMBER: US 60/187,470
;; PRIOR FILING DATE: 2000-03-06
;; NUMBER OF SEQ ID NOS: 482
;; SOFTWARE: Patent.pm
;; SEQ ID NO 357
;; LENGTH: 183
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: SIGNAL
;; LOCATION: -47...-1
US-09-731-872-357

Query Match
Best Local Similarity 100.0%; Score 21; DB 10; Length 183;
Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
Db 143 HARL 146

RESULT 13
US-09-815-242-11908
;; Sequence 11908, Application US/09815242
;; Patent No. US20020061569A1
;; GENERAL INFORMATION:
;; APPLICANT: Haselbeck, Robert
;; APPLICANT: Ohlsen, Karl L.
;; APPLICANT: Zyskind, Judith W.
;; APPLICANT: Wall, Daniel
;; APPLICANT: Trawick, John D.
;; APPLICANT: Carr, Grant J.
;; APPLICANT: Yamamoto, Robert T.
;; TITLE OF INVENTION: Identification of Essential Genes in
;; FILE REFERENCE: ELITRA.011A
;; CURRENT APPLICATION NUMBER: US/09/815,242
;; PRIOR FILING DATE: 2001-03-21
;; PRIOR APPLICATION NUMBER: 60/191,078
;; PRIOR FILING DATE: 2000-03-21
;; PRIOR APPLICATION NUMBER: 60/206,848
;; PRIOR FILING DATE: 2000-05-23

;; PRIOR APPLICATION NUMBER: 60/207,727
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: 60/242,578
;; PRIOR FILING DATE: 2000-10-23
;; PRIOR APPLICATION NUMBER: 60/253,625
;; PRIOR FILING DATE: 2000-11-27
;; PRIOR APPLICATION NUMBER: 60/257,931
;; PRIOR FILING DATE: 2000-12-22
;; PRIOR APPLICATION NUMBER: 60/269,308
;; PRIOR FILING DATE: 2001-02-16
;; NUMBER OF SEQ ID NOS: 14110
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 11908
;; LENGTH: 186
;; TYPE: PRT
;; ORGANISM: Pseudomonas aeruginosa
US-09-815-242-11908

Query Match 100.0%; Score 21; DB 10; Length 186;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
|||
Db 13 HARL 16

RESULT 14
US-09-682-667-2
;; Sequence 2, Application US/09682667
;; Patent No. US20020076794A1
;; GENERAL INFORMATION:
;; APPLICANT: Hsueh, Aaron J.W.
;; APPLICANT: Hsu, Sheau Yu
;; TITLE OF INVENTION: MAMMALIAN PRO-APOPTOTIC BOK GENES AND
;; TITLE OF INVENTION: THEIR USES
;; FILE REFERENCE: STAN072CON
;; CURRENT APPLICATION NUMBER: US/09/682,667
;; CURRENT FILING DATE: 2001-10-05
;; PRIOR APPLICATION NUMBER: 09/517,358
;; PRIOR FILING DATE: 2000-03-02
;; PRIOR APPLICATION NUMBER: 09/186,250
;; PRIOR FILING DATE: 1998-11-04
;; PRIOR APPLICATION NUMBER: 60/064,943
;; PRIOR FILING DATE: 1997-11-07
;; NUMBER OF SEQ ID NOS: 18
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 2
;; LENGTH: 213
;; TYPE: PRT
;; ORGANISM: R. rattus
US-09-682-667-2

Query Match 100.0%; Score 21; DB 10; Length 213;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
|||
Db 40 HARL 43

RESULT 15
US-09-682-667-6
;; Sequence 6, Application US/09682667
;; Patent No. US20020076794A1
;; GENERAL INFORMATION:
;; APPLICANT: Hsueh, Aaron J.W.
;; APPLICANT: Hsu, Sheau Yu
;; TITLE OF INVENTION: MAMMALIAN PRO-APOPTOTIC BOK GENES AND
;; TITLE OF INVENTION: THEIR USES
;; FILE REFERENCE: STAN072CON
;; CURRENT APPLICATION NUMBER: US/09/682,667

;; CURRENT FILING DATE: 2001-10-05
;; PRIOR APPLICATION NUMBER: 09/517,358
;; PRIOR FILING DATE: 2000-03-02
;; PRIOR APPLICATION NUMBER: 09/186,250
;; PRIOR FILING DATE: 1998-11-04
;; PRIOR APPLICATION NUMBER: 60/064,943
;; PRIOR FILING DATE: 1997-11-07
;; NUMBER OF SEQ ID NOS: 18
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 6
;; LENGTH: 213
;; TYPE: PRT
;; ORGANISM: H.sapiens
US-09-682-667-6

Query Match 100.0%; Score 21; DB 10; Length 213;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
|||
Db 40 HARL 43

RESULT 16
US-09-750-373-22
;; Sequence 22, Application US/09750373
;; Patent No. US20020062013A1
;; GENERAL INFORMATION:
;; APPLICANT: Lind, Peter
;; APPLICANT: Wood, Linda S.
;; APPLICANT: Hiesch, Ronald
;; APPLICANT: Ruff, Valerie
;; APPLICANT: Lindberg, Eleni
;; APPLICANT: Parodi, Luis A.
;; APPLICANT: Vogeli, Gabriel
;; TITLE OF INVENTION: No. US20020062013A1el G Protein Coupled Receptors
;; FILE REFERENCE: PHRM-0300
;; CURRENT APPLICATION NUMBER: US/09/750,373
;; CURRENT FILING DATE: 2000-12-28
;; PRIOR APPLICATION NUMBER: 60/184,305
;; PRIOR FILING DATE: 2000-02-23
;; PRIOR APPLICATION NUMBER: 60/188,880
;; PRIOR FILING DATE: 2000-03-13
;; PRIOR APPLICATION NUMBER: 60/219,492
;; PRIOR FILING DATE: 2000-07-20
;; PRIOR APPLICATION NUMBER: 60/173,339
;; PRIOR FILING DATE: 1999-12-28
;; PRIOR APPLICATION NUMBER: 60/224,321
;; PRIOR FILING DATE: 2000-08-11
;; PRIOR APPLICATION NUMBER: 60/200,534
;; PRIOR FILING DATE: 2000-04-27
;; PRIOR APPLICATION NUMBER: 60/239,062
;; PRIOR FILING DATE: 2000-10-09
;; NUMBER OF SEQ ID NOS: 56
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO 22
;; LENGTH: 216
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-750-373-22

Query Match 100.0%; Score 21; DB 10; Length 216;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
|||
Db 70 HARL 73

RESULT 17
US-09-976-736-14

```
Sequence 14, Application US/09976736
Patent No. US20020161178A1
GENERAL INFORMATION:
APPLICANT: Bass, Michael B
APPLICANT: Sullivan, John K
APPLICANT: Theill, Lars E
APPLICANT: Wang, Daquang
TITLE OF INVENTION: NOVEL DKR POLYPEPTIDES
FILE REFERENCE: A-548
CURRENT APPLICATION NUMBER: US/09/976,736
PRIOR FILING DATE: 2001-10-09
PRIOR APPLICATION NUMBER: US/09/161,241
NUMBER OF SEQ ID NOS: 78
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 14
LENGTH: 224
TYPE: PRT
ORGANISM: Human
US-09-976-736-14

Query Match
Best Local Similarity 100.0%; Score 21; DB 9; Length 224;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
Db 212 HARL 215

RESULT 18
US-09-323-998D-15
Sequence 15, Application US/09323998D
Patent No. US20020102631A1
GENERAL INFORMATION:
APPLICANT: CUNNINGHAM JR., FRANCIS X.
TITLE OF INVENTION: GENES OF CAROTENOID BIOSYNTHESIS AND METABOLISM AND
FILE REFERENCE: 108172-09019
CURRENT APPLICATION NUMBER: US/09/323,998D
CURRENT FILING DATE: 1999-06-02
PRIOR APPLICATION NUMBER: 09/088,724
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 09/088,725
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 08/937,155
PRIOR FILING DATE: 1997-09-25
PRIOR APPLICATION NUMBER: 08/624,125
PRIOR FILING DATE: 1996-03-29
NUMBER OF SEQ ID NOS: 61
SOFTWARE: Patentln Ver. 2.1
SEQ ID NO 15
LENGTH: 293
TYPE: PRT
ORGANISM: Haematococcus pluvialis
US-09-323-998D-15

Query Match
Best Local Similarity 100.0%; Score 21; DB 10; Length 293;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
Db 25 HARL 28

RESULT 19
US-09-323-998D-43
Sequence 43, Application US/09323998D
Patent No. US20020102631A1
GENERAL INFORMATION:
APPLICANT: CUNNINGHAM JR., FRANCIS X.
```

```
APPLICANT: SUN, ZAIREN
TITLE OF INVENTION: GENES OF CAROTENOID BIOSYNTHESIS AND METABOLISM AND
FILE REFERENCE: 108172-09019
CURRENT APPLICATION NUMBER: US/09/323,998D
CURRENT FILING DATE: 1999-06-02
PRIOR APPLICATION NUMBER: 09/088,724
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 09/088,725
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 08/937,155
PRIOR FILING DATE: 1997-09-25
PRIOR APPLICATION NUMBER: 08/624,125
PRIOR FILING DATE: 1996-03-29
NUMBER OF SEQ ID NOS: 61
SOFTWARE: Patentln Ver. 2.1
SEQ ID NO 43
LENGTH: 293
TYPE: PRT
ORGANISM: Haematococcus pluvialis
US-09-323-998D-43

Query Match
Best Local Similarity 100.0%; Score 21; DB 10; Length 293;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
Db 25 HARL 28

RESULT 20
US-09-815-242-13325
Sequence 13325, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13325
LENGTH: 298
TYPE: PRT
ORGANISM: Streptococcus pneumoniae
US-09-815-242-13325

Query Match
Best Local Similarity 100.0%; Score 21; DB 10; Length 298;
```

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
|||||

Db 147 HARL 150

RESULT 21

US-09-815-242-13603
; Sequence 13603, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlson, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13603
; LENGTH: 298
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-815-242-13603

Query Match 100.0%; Score 21; DB 10; Length 298;

Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
|||||

Db 147 HARL 150

RESULT 22

US-09-323-998D-14
; Sequence 14, Application US/09323998D
; Patent No. US20020102631A1
; GENERAL INFORMATION:
; APPLICANT: CUNNINGHAM JR., FRANCIS X.
; APPLICANT: SUN, ZAIREN
; TITLE OF INVENTION: GENES OF CAROTENOID BIOSYNTHESIS AND METABOLISM AND
; FILE REFERENCE: 108172-09019
; CURRENT APPLICATION NUMBER: US/09/323,998D
; PRIOR FILING DATE: 1999-06-02
; PRIOR APPLICATION NUMBER: 09/088,724
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 09/088,725
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 08/937,155

; PRIOR FILING DATE: 1997-09-25
; PRIOR APPLICATION NUMBER: 08/624,125
; PRIOR FILING DATE: 1996-03-29
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 14
; LENGTH: 305
; TYPE: PRT
; ORGANISM: Haematococcus pluvialis
US-09-323-998D-14

US-09-323-998D-14

Query Match 100.0%; Score 21; DB 10; Length 305;

Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
|||||

Db 25 HARL 28

RESULT 23

US-09-323-998D-44
; Sequence 44, Application US/09323998D
; Patent No. US20020102631A1
; GENERAL INFORMATION:
; APPLICANT: CUNNINGHAM JR., FRANCIS X.
; APPLICANT: SUN, ZAIREN
; TITLE OF INVENTION: GENES OF CAROTENOID BIOSYNTHESIS AND METABOLISM AND
; FILE REFERENCE: 108172-09019
; CURRENT APPLICATION NUMBER: US/09/323,998D
; PRIOR FILING DATE: 1999-06-02
; PRIOR APPLICATION NUMBER: 09/088,724
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 09/088,725
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 08/937,155
; PRIOR FILING DATE: 1997-09-25
; PRIOR APPLICATION NUMBER: 08/624,125
; PRIOR FILING DATE: 1996-03-29
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 44
; LENGTH: 305
; TYPE: PRT
; ORGANISM: Haematococcus pluvialis
US-09-323-998D-44

Query Match 100.0%; Score 21; DB 10; Length 305;

Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
|||||

Db 25 HARL 28

RESULT 24

US-09-815-242-4904
; Sequence 4904, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlson, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A

```
;; CURRENT APPLICATION NUMBER: US/09/815,242
;; CURRENT FILING DATE: 2001-03-21
;; PRIOR APPLICATION NUMBER: 60/191,078
;; PRIOR FILING DATE: 2000-03-21
;; PRIOR APPLICATION NUMBER: 60/206,848
;; PRIOR FILING DATE: 2000-05-23
;; PRIOR APPLICATION NUMBER: 60/207,727
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: 60/242,578
;; PRIOR FILING DATE: 2000-10-23
;; PRIOR APPLICATION NUMBER: 60/253,625
;; PRIOR FILING DATE: 2000-11-27
;; PRIOR APPLICATION NUMBER: 60/257,931
;; PRIOR FILING DATE: 2000-12-22
;; PRIOR APPLICATION NUMBER: 60/269,308
;; PRIOR FILING DATE: 2001-02-16
;; NUMBER OF SEQ ID NOS: 14110
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 4904
;; LENGTH: 312
;; TYPE: PRY
;; ORGANISM: Enterococcus faecalis
US-09-815-242-4904
```

```
Query Match          100.0%; Score 21; DB 10; Length 312;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 HARL 4
        ||||
Db      135 HARL 138
```

```
RESULT 25
US-09-815-242-10541
;; Sequence 10541, Application US/09815242
;; Patent No. US20020061569A1
;; GENERAL INFORMATION:
;; APPLICANT: Haselbeck, Robert
;; APPLICANT: Ohlsen, Karl L.
;; APPLICANT: Zyskind, Judith W.
;; APPLICANT: Wall, Daniel
;; APPLICANT: Trawick, John D.
;; APPLICANT: Carr, Grant J.
;; APPLICANT: Yamamoto, Robert T.
;; APPLICANT: Xu, H. Howard
;; TITLE OF INVENTION: Identification of Essential Genes in
;; FILE REFERENCE: ELITRA.011A
;; CURRENT APPLICATION NUMBER: US/09/815,242
;; CURRENT FILING DATE: 2001-03-21
;; PRIOR APPLICATION NUMBER: 60/191,078
;; PRIOR FILING DATE: 2000-03-21
;; PRIOR APPLICATION NUMBER: 60/206,848
;; PRIOR FILING DATE: 2000-05-23
;; PRIOR APPLICATION NUMBER: 60/207,727
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: 60/242,578
;; PRIOR FILING DATE: 2000-10-23
;; PRIOR APPLICATION NUMBER: 60/253,625
;; PRIOR FILING DATE: 2000-11-27
;; PRIOR APPLICATION NUMBER: 60/257,931
;; PRIOR FILING DATE: 2000-12-22
;; PRIOR APPLICATION NUMBER: 60/269,308
;; PRIOR FILING DATE: 2001-02-16
;; NUMBER OF SEQ ID NOS: 14110
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 10541
;; LENGTH: 312
;; TYPE: PRY
;; ORGANISM: Enterococcus faecalis
US-09-815-242-10541
```

```
Query Match          100.0%; Score 21; DB 10; Length 312;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 HARL 4
        ||||
Db      135 HARL 138
```

```
RESULT 26
US-09-815-242-11761
;; Sequence 11761, Application US/09815242
;; Patent No. US20020061569A1
;; GENERAL INFORMATION:
;; APPLICANT: Haselbeck, Robert
;; APPLICANT: Ohlsen, Karl L.
;; APPLICANT: Zyskind, Judith W.
;; APPLICANT: Wall, Daniel
;; APPLICANT: Trawick, John D.
;; APPLICANT: Carr, Grant J.
;; APPLICANT: Yamamoto, Robert T.
;; APPLICANT: Xu, H. Howard
;; TITLE OF INVENTION: Identification of Essential Genes in
;; FILE REFERENCE: ELITRA.011A
;; CURRENT APPLICATION NUMBER: US/09/815,242
;; CURRENT FILING DATE: 2001-03-21
;; PRIOR APPLICATION NUMBER: 60/191,078
;; PRIOR FILING DATE: 2000-03-21
;; PRIOR APPLICATION NUMBER: 60/206,848
;; PRIOR FILING DATE: 2000-05-23
;; PRIOR APPLICATION NUMBER: 60/207,727
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: 60/242,578
;; PRIOR FILING DATE: 2000-10-23
;; PRIOR APPLICATION NUMBER: 60/253,625
;; PRIOR FILING DATE: 2000-11-27
;; PRIOR APPLICATION NUMBER: 60/257,931
;; PRIOR FILING DATE: 2000-12-22
;; PRIOR APPLICATION NUMBER: 60/269,308
;; PRIOR FILING DATE: 2001-02-16
;; NUMBER OF SEQ ID NOS: 14110
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 11761
;; LENGTH: 331
;; TYPE: PRY
;; ORGANISM: Klebsiella pneumoniae
US-09-815-242-11761
```

```
Query Match          100.0%; Score 21; DB 10; Length 331;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 HARL 4
        ||||
Db      228 HARL 231
```

```
RESULT 27
US-09-361-655-14
;; Sequence 14, Application US/09361655
;; Patent No. US20020161204A1
;; GENERAL INFORMATION:
;; APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE
;; APPLICANT: LEE, Se-Jin
;; APPLICANT: ESQUELA, Aurora F.
;; TITLE OF INVENTION: METHODS OF DETECTING LIVER CELLS EXPRESSING GROWTH DIFFERENTI
;; FILE REFERENCE: JH01220-4
;; CURRENT APPLICATION NUMBER: US/09/361,655
;; CURRENT FILING DATE: 1999-07-27
;; PRIOR APPLICATION NUMBER: US 08/765,662
;; PRIOR FILING DATE: 1997-04-28
```

```

; PRIOR APPLICATION NUMBER: PCT/ US95/08745
; PRIOR FILING DATE: 1995-07-12
; PRIOR APPLICATION NUMBER: US 08/274,215
; PRIOR FILING DATE: 1994-07-13
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patent version 3.1
; SEQ ID NO 14
; LENGTH: 350
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-361-655-14

Query Match
Best Local Similarity 100.0%; Score 21; DB 9; Length 350;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HARL 4
Db 123 HARL 126

RESULT 28
US-09-860-351-2
; Sequence 2, Application US/09860351
; Patent No. US20020077463A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Meyers, Rachel
; TITLE OF INVENTION: 16105, A NOVEL PROTEIN HUMAN PHOSPHATASE
; FILE REFERENCE: 38155-20013.00
; CURRENT APPLICATION NUMBER: US/09/860,351
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: US 60/205,260
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-860-351-2

Query Match
Best Local Similarity 100.0%; Score 21; DB 10; Length 352;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HARL 4
Db 174 HARL 177

RESULT 29
US-09-815-242-12071
; Sequence 12071, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
```

```

; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12071
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-815-242-12071

Query Match
Best Local Similarity 100.0%; Score 21; DB 10; Length 358;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HARL 4
Db 142 HARL 145

RESULT 30
US-09-815-242-5116
; Sequence 5116, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5116
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-815-242-5116

Query Match
Best Local Similarity 100.0%; Score 21; DB 10; Length 359;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HARL 4
```

Db 276 HARL 279

|||||
RESULT 31
US-09-833-790-253
; Sequence 253, Application US/09833790
; Patent No. US20020068288A1
; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Wang, Tongtong
; APPLICANT: Secrist, Heather
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Indrias, Carol Y.
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.512
; CURRENT APPLICATION NUMBER: US/09/833,790
; CURRENT FILING DATE: 2001-04-11
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 253
; LENGTH: 370
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-833-790-253

Query Match 100.0%; Score 21; DB 10; Length 370;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
|||||
Db 311 HARL 314

RESULT 32
US-09-973-941-4
; Sequence 4, Application US/09973941
; Patent No. US2002016455A1
; GENERAL INFORMATION:
; APPLICANT: Koch, Jean-Marc
; APPLICANT: Bartel, Paul L.
; APPLICANT: Heichman, Karen
; TITLE OF INVENTION: Protein-Protein Interactions in Neurodegenerative
; FILE REFERENCE: Diseases
; CURRENT APPLICATION NUMBER: US/09/973,941
; CURRENT FILING DATE: 2001-10-11
; PRIOR APPLICATION NUMBER: US 60/240,790
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: US 60/304,775
; PRIOR FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-973-941-4

Query Match 100.0%; Score 21; DB 9; Length 372;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
|||||
Db 174 HARL 177

RESULT 33
US-09-973-963-4

;; Sequence 4, Application US/09973963
;; Patent No. US20020106676A1
;; GENERAL INFORMATION:
;; APPLICANT: Koch, Jean-Marc
;; APPLICANT: Bartel, Paul L.
;; APPLICANT: Heichman, Karen
;; TITLE OF INVENTION: Protein-Protein Interactions in Neurodegenerative
;; FILE REFERENCE: Diseases
;; CURRENT APPLICATION NUMBER: US/09/973,963
;; CURRENT FILING DATE: 2001-10-11
;; PRIOR APPLICATION NUMBER: US 60/240,790
;; PRIOR FILING DATE: 2000-10-17
;; PRIOR APPLICATION NUMBER: US 60/304,775
;; PRIOR FILING DATE: 2001-07-13
;; NUMBER OF SEQ ID NOS: 8
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 4
;; LENGTH: 372
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-973-963-4

Query Match 100.0%; Score 21; DB 10; Length 372;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
|||||
Db 174 HARL 177

RESULT 34
US-09-973-064-4
; Sequence 4, Application US/09973064
; Patent No. US20020106773A1
; GENERAL INFORMATION:
; APPLICANT: Koch, Jean-Marc
; APPLICANT: Bartel, Paul L.
; APPLICANT: Heichman, Karen
; TITLE OF INVENTION: Protein-Protein Interactions in Neurodegenerative
; FILE REFERENCE: Diseases
; CURRENT APPLICATION NUMBER: US/09/973,064
; CURRENT FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/240,790
; PRIOR FILING DATE: 2000-10-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-973-064-4

Query Match 100.0%; Score 21; DB 10; Length 372;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
|||||
Db 174 HARL 177

RESULT 35
US-09-973-077-4
; Sequence 4, Application US/09973077
; Patent No. US2002011479A1
; GENERAL INFORMATION:
; APPLICANT: Koch, Jean-Marc
; APPLICANT: Bartel, Paul L.
; APPLICANT: Heichman, Karen
; TITLE OF INVENTION: Protein-Protein Interactions in Neurodegenerative


```

; TITLE OF INVENTION: Diseases
; FILE REFERENCE: Protein Interactions in ND
; CURRENT APPLICATION NUMBER: US/09/973,077
; CURRENT FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/240,790
; PRIOR FILING DATE: 2000-10-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 4
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-973-077-4

```

```

Query Match      100.0%; Score 21; DB 10; Length 372;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```
OY      1 HARL 4
        ||||
Db       174 HARL 177

```

```

RESULT 36
US-09-973-063-4
; Sequence 4, Application US/09973063
; Patent No. US2002011519A1
; GENERAL INFORMATION:
; APPLICANT: Roch, Jean-Marc
; APPLICANT: Bartel, Paul L.
; APPLICANT: Heichman, Karen
; TITLE OF INVENTION: Protein-Protein Interactions in Neurodegenerative
; FILE REFERENCE: Protein Interactions in ND
; CURRENT APPLICATION NUMBER: US/09/973,063
; CURRENT FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/240,790
; PRIOR FILING DATE: 2000-10-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 4
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-973-063-4

```

```

Query Match      100.0%; Score 21; DB 10; Length 372;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```
OY      1 HARL 4
        ||||
Db       174 HARL 177

```

```

RESULT 37
US-09-973-964-4
; Sequence 4, Application US/09973964
; Patent No. US20020115606A1
; GENERAL INFORMATION:
; APPLICANT: Roch, Jean-Marc
; APPLICANT: Bartel, Paul L.
; APPLICANT: Heichman, Karen
; TITLE OF INVENTION: Protein-Protein Interactions in Neurodegenerative
; FILE REFERENCE: Protein Interactions in ND
; CURRENT APPLICATION NUMBER: US/09/973,964
; CURRENT FILING DATE: 2001-10-11
; PRIOR APPLICATION NUMBER: US 60/240,790
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: US 60/304,775
; PRIOR FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 8

```

```

; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 4
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-973-964-4

```

```

Query Match      100.0%; Score 21; DB 10; Length 372;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```
OY      1 HARL 4
        ||||
Db       174 HARL 177

```

```

RESULT 38
US-09-975-072-4
; Sequence 4, Application US/09975072
; Patent No. US20020115607A1
; GENERAL INFORMATION:
; APPLICANT: Roch, Jean-Marc
; APPLICANT: Bartel, Paul L.
; APPLICANT: Heichman, Karen
; TITLE OF INVENTION: Protein-Protein Interactions in Neurodegenerative
; FILE REFERENCE: Protein Interactions in ND
; CURRENT APPLICATION NUMBER: US/09/975,072
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,790
; PRIOR FILING DATE: 2000-10-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 4
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-975-072-4

```

```

Query Match      100.0%; Score 21; DB 10; Length 372;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```
OY      1 HARL 4
        ||||
Db       174 HARL 177

```

```

RESULT 39
US-09-972-038-4
; Sequence 4, Application US/09972038
; Patent No. US20020119155A1
; GENERAL INFORMATION:
; APPLICANT: Roch, Jean-Marc
; APPLICANT: Bartel, Paul L.
; APPLICANT: Heichman, Karen
; TITLE OF INVENTION: Protein-Protein Interactions in Neurodegenerative
; FILE REFERENCE: Protein Interactions in ND
; CURRENT APPLICATION NUMBER: US/09/972,038
; CURRENT FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: US 60/240,790
; PRIOR FILING DATE: 2000-10-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 4
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-972-038-4

```

```

Query Match      100.0%; Score 21; DB 10; Length 372;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;

```

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HARL 4
1111

DB 174 HARL 177

RESULT 40

US-09-972-757-4

; Sequence 4, Application US/09972757
; Patent No. US20020119927A1

; GENERAL INFORMATION:

; APPLICANT: Roch, Jean-Marc
; APPLICANT: Bartel, Paul L.

; TITLE OF INVENTION: Protein-Protein Interactions in Neurodegenerative

; FILE REFERENCE: Protein Interactions in ND

; CURRENT APPLICATION NUMBER: US/09/972,757

; PRIOR FILING DATE: 2001-10-09

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 4

; LENGTH: 372

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-972-757-4

Query Match 100.0%; Score 21; DB 10; Length 372;

Best Local Similarity 100.0%; Pred. No. 2.7e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HARL 4
1111

DB 174 HARL 177

RESULT 41

US-09-973-965-4

; Sequence 4, Application US/09973965
; Patent No. US20020124273A1

; GENERAL INFORMATION:

; APPLICANT: Roch, Jean-Marc
; APPLICANT: Bartel, Paul L.

; TITLE OF INVENTION: Protein-Protein Interactions in Neurodegenerative

; FILE REFERENCE: Protein Interactions in ND

; CURRENT APPLICATION NUMBER: US/09/973,965

; PRIOR FILING DATE: 2001-10-11

; NUMBER OF SEQ ID NOS: 8

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 4

; LENGTH: 372

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-973-965-4

Query Match 100.0%; Score 21; DB 10; Length 372;

Best Local Similarity 100.0%; Pred. No. 2.7e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HARL 4
1111

DB 174 HARL 177

RESULT 42

US-09-925-300-1655

; Sequence 1655, Application US/09925300
; Patent No. US20020151681A1

; GENERAL INFORMATION:

; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben,

; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies

; FILE REFERENCE: PA101

; CURRENT APPLICATION NUMBER: US/09/925,300

; PRIOR FILING DATE: 2001-08-10

; NUMBER OF SEQ ID NOS: 1890

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 1655

; LENGTH: 373

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-925-300-1655

Query Match 100.0%; Score 21; DB 10; Length 373;

Best Local Similarity 100.0%; Pred. No. 2.7e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HARL 4
1111

DB 175 HARL 178

RESULT 43

US-09-964-666-2

; Sequence 2, Application US/09964666
; Patent No. US20020104108A1

; GENERAL INFORMATION:

; APPLICANT: de la Monte, Suzanne
; APPLICANT: Rands, Jack R.

; TITLE OF INVENTION: Transgenic Animals and Cell Lines for

; Screening Drugs Effective for the Treatment or Prevention

; of Alzheimer's Disease

; NUMBER OF SEQUENCE ADDRESSES: 14

; CORRESPONDENCE ADDRESS:

; ADDRESSSEE: Sterne, Kessler, Goldstein & Fox, P.L.L.C.

; STREET: 1100 New York Ave., Suite 600

; CITY: Washington

; STATE: DC

; COUNTRY: USA

; ZIP: 20005-3934

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA: US/09/964,666

; APPLICATION NUMBER: US/09/964,666

; FILING DATE: 28-Sep-2001

; CLASSIFICATION: <Unknown>

; ATTORNEY/AGENT INFORMATION:

NAME: Esmond, Robert W.
REGISTRATION NUMBER: 32,893
REFERENCE/DOCKET NUMBER: 0609.4370000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 375 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-964-666-2

Query Match 100.0%; Score 21; DB 10; Length 375;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
Db 292 HARL 295

RESULT 44
US-09-964-412-2
Sequence 2, Application US/09964412
Patent No. US20020129391A1
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
Mands, Jack R.
TITLE OF INVENTION: Transgenic Animals and Cell Lines for
Screening Drugs Effective for the Treatment or Prevention
of Alzheimer's Disease
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRES:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox, P.L.L.C.
STREET: 1100 New York Ave., Suite 600
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/964,412
FILING DATE: 28-Sep-2001
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Esmond, Robert W.
REGISTRATION NUMBER: 32,893
REFERENCE/DOCKET NUMBER: 0609.4370000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 375 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-964-412-2

Query Match 100.0%; Score 21; DB 10; Length 375;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
Db 292 HARL 295

Db 292 HARL 295

RESULT 45
US-09-371-900-38
Sequence 38, Application US/09371900
Patent No. US20020137700A1
GENERAL INFORMATION:
APPLICANT: FALB, DEAN A
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
TREATMENT AND DIAGNOSIS OF CARDIOVASCULAR DISEASE
NUMBER OF SEQUENCES: 54
CORRESPONDENCE ADDRESS:
ADDRESSEE: PENNIE & EDMONDS LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/371,900
FILING DATE: 11-Aug-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/599,654
FILING DATE: 09-FEB-1996
APPLICATION NUMBER: US 08/485,573
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/386,844
FILING DATE: 10-FEB-1995
ATTORNEY/AGENT INFORMATION:
NAME: CORUZZI, LAURA A
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7853-104
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 375 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: unknown
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 38:
US-09-371-900-38

Query Match 100.0%; Score 21; DB 10; Length 375;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
Db 173 HARL 176

RESULT 46
US-09-924-417-59
Sequence 59, Application US/09924417
Patent No. US20020142441A1
GENERAL INFORMATION:
APPLICANT: Falb, Dean
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
THE TREATMENT AND DIAGNOSIS OF CARDIOVASCULAR
DISEASE
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:

ADDRESSEE: PENNIE & EDMONDS LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA: US/09/924,417
APPLICATION NUMBER: US/09/924,417
FILING DATE: 07-Aug-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/034,286
FILING DATE: 04-MAR-1998
APPLICATION NUMBER: 08/870,434
FILING DATE: 06-JUN-1997
APPLICATION NUMBER: 08/799,910
FILING DATE: 13-FEB-1997
APPLICATION NUMBER: 60/011,787
FILING DATE: 16-FEB-1996
APPLICATION NUMBER: 08/599,654
FILING DATE: 09-FEB-1996
APPLICATION NUMBER: 08/485,573
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: 08/386,844
FILING DATE: 10-FEB-1995
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7853-114-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)7909090
TELEFAX: (212)8699741
INFORMATION FOR SEQ ID NO: 59:
SEQUENCE CHARACTERISTICS:
LENGTH: 375 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 59:
US-09-924-417-59
Query Match 100.0%; Score 21; DB 10; Length 375;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HARL 4
DB 173 HARL 176
RESULT 47
US-09-833-745-9
Sequence 9, Application US/09833745
Patent No. US20020052038A1
GENERAL INFORMATION:
APPLICANT: ROBERTS, JOSEPH
APPLICANT: SETHURAMAN, NATARAJAN
APPLICANT: MACALLISTER, THOMAS
TITLE OF INVENTION: CLONING, OVEREXPRESSION AND THERAPEUTIC USE OF
FILE REFERENCE: 078728/0106
CURRENT APPLICATION NUMBER: US/09/833,745
CURRENT FILING DATE: 2001-04-13
PRIOR APPLICATION NUMBER: 60/197,770
PRIOR FILING DATE: 2000-04-14
NUMBER OF SEQ ID NOS: 66

SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 9
LENGTH: 404
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-833-745-9
Query Match 100.0%; Score 21; DB 10; Length 404;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HARL 4
DB 303 HARL 306
RESULT 48
US-09-833-745-4
Sequence 4, Application US/09833745
Patent No. US20020052038A1
GENERAL INFORMATION:
APPLICANT: ROBERTS, JOSEPH
APPLICANT: SETHURAMAN, NATARAJAN
APPLICANT: MACALLISTER, THOMAS
TITLE OF INVENTION: CLONING, OVEREXPRESSION AND THERAPEUTIC USE OF
FILE REFERENCE: 078728/0106
CURRENT APPLICATION NUMBER: US/09/833,745
CURRENT FILING DATE: 2001-04-13
PRIOR APPLICATION NUMBER: 60/197,770
PRIOR FILING DATE: 2000-04-14
NUMBER OF SEQ ID NOS: 66
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 405
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-833-745-4
Query Match 100.0%; Score 21; DB 10; Length 405;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HARL 4
DB 304 HARL 307
RESULT 49
US-09-712-363-173
Sequence 173, Application US/09712363
Patent No. US20020164588A1
GENERAL INFORMATION:
APPLICANT: Eisenberg, David
APPLICANT: Rotstein, Sergio H.
APPLICANT: Marcotte, Edward M.
TITLE OF INVENTION: DETERMINING THE FUNCTIONS AND
FILE REFERENCE: 07419-032001
CURRENT APPLICATION NUMBER: US/09/712,363
CURRENT FILING DATE: 2000-11-13
PRIOR APPLICATION NUMBER: PCT/US00/02246
PRIOR FILING DATE: 2000-01-28
PRIOR APPLICATION NUMBER: 60/179,531
PRIOR FILING DATE: 2000-02-01
PRIOR APPLICATION NUMBER: 60/117,844
PRIOR FILING DATE: 1999-01-29

; PRIOR APPLICATION NUMBER: 60/118,206,
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: 60/126,593
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 60/134,093
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 60/134,092
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 60/165,124
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/165,086
; PRIOR FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 292
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 173
; LENGTH: 410
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-712-363-173

Query Match 100.0%; Score 21; DB 9; Length 410;
Best Local Similarity 100.0%; Pred. No. 3e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
|||
Db 129 HARL 132

RESULT 50
US-09-925-301-954
; Sequence 954, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 954
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-301-954

Query Match 100.0%; Score 21; DB 10; Length 428;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
|||
Db 165 HARL 168

Search completed: November 13, 2002, 13:45:32
Job time : 11 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2002, 13:22:18 ; Search time 35 Seconds

(without alignments)
15.229 Million cell updates/sec

Title: US-09-697-590-2_COPY_292_295

Perfect score: 21

Sequence: 1 HARL 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

1: /SID2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SID2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SID2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SID2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SID2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SID2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SID2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SID2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SID2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SID2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SID2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SID2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SID2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SID2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SID2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SID2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SID2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SID2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SID2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SID2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SID2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SID2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	21	100.0	4	23	ABB81512
2	21	100.0	5	23	ABB81511
3	21	100.0	5	23	ABB81513
4	21	100.0	6	23	ABB81514
5	21	100.0	6	23	ABB81521
6	21	100.0	6	23	ABB81528
7	21	100.0	7	23	ABB81515
8	21	100.0	7	23	ABB81517
9	21	100.0	7	23	ABB81518
10	21	100.0	8	23	ABB81537

11	21	100.0	9	22	AAG39428
12	21	100.0	9	22	AAG39478
13	21	100.0	10	22	AAG94522
14	21	100.0	10	23	ABB81536
15	21	100.0	12	22	AAB31270
16	21	100.0	12	23	ABB81532
17	21	100.0	13	23	ABB81532
18	21	100.0	14	22	AAK00412
19	21	100.0	14	22	AAK00444
20	21	100.0	14	23	ABB81534
21	21	100.0	15	21	AAV79924
22	21	100.0	15	23	ABB81535
23	21	100.0	16	23	ABB81530
24	21	100.0	17	16	AAK80293
25	21	100.0	25	21	AAV58899
26	21	100.0	30	20	AAW94846
27	21	100.0	30	21	AAV56381
28	21	100.0	30	23	AAU85095
29	21	100.0	30	23	AAU85096
30	21	100.0	31	20	AAW94447
31	21	100.0	32	21	ABB34676
32	21	100.0	32	22	ABB31475
33	21	100.0	32	22	ABB36687
34	21	100.0	32	22	ABB22024
35	21	100.0	32	22	AAW57451
36	21	100.0	32	22	AAW69846
37	21	100.0	32	22	AAW17668
38	21	100.0	32	22	AAW30185
39	21	100.0	32	22	AAW05331
40	21	100.0	32	23	ABG39474
41	21	100.0	36	22	AAO09129
42	21	100.0	37	22	AAO08247
43	21	100.0	40	22	AAW94835
44	21	100.0	42	22	AAO06502
45	21	100.0	43	22	AAU20929

ALIGNMENTS

RESULT 1
ABB81512 standard; peptide; 4 AA.
XX ABB81512:
XX
XX 02-SEP-2002 (first entry)
XX
XX Neural thread protein (NTP) peptide #2.
DE
XX
KW Neural thread protein; NTP; Harl1 peptide; Alzheimer's disease;
KW Down's syndrome; neuroectodermal tumour; astrocytoma; glioblastoma;
KW hypoxia; ischemia; cerebral infarction.
XX
XX Homo sapiens.
OS
XX
XX W0200234915-A2.
PN
XX
XX 02-MAY-2002.
PD
XX
XX 25-OCT-2001; 2001WO-US42813.
PF
XX
XX 27-OCT-2000; 2000US-0697590.
PR
XX
XX (NYMO-) NYMOX PHARM CORP.
PA
XX
XX Fitzpatrick J, Averbach P, Focht MSS, Bibiano R;
PI
XX
XX WPI; 2002-507998/54.
DR
XX
XX New Harl1 peptide sequences of the Neural Thread protein, useful in
PT therapeutic assays, e.g. as targets for developing drugs for treating
PT Alzheimer's disease, ischemia or cerebral infarction, or for diagnosing

Pt		these diseases -
Xx		
Ps	Claim 1; Page 29; 53pp; English.	
Xx		
Cc	The present invention describes a neural thread protein (NTP) peptide	
Cc	having an amino acid sequence selected from ABB81511 to ABB81529 and	
Cc	their homologues, which are referred collectively as Harill peptides (I).	
Cc	(I) have neuroprotective, nootropic, vasotropic and cerebroprotective	
Cc	activities, and can be used in peptide therapy. The Harill peptide	
Cc	sequences can be used as analogues for NTP in therapeutic or diagnostic	
Cc	assays by replacing NTP with the peptide in such an assay. The Harill	
Cc	peptides are also useful as a trap material in a diagnostic or	
Cc	therapeutic assay. Therefore, the Harill peptides are useful in binding	
Cc	assays, protein and antibody purification, therapeutics or diagnostics.	
Cc	In particular, the peptides are also useful for diagnosing Alzheimer's	
Cc	disease, Down's syndrome, neuroectodermal tumours, astrocytomas,	
Cc	glioblastomas, hypoxia, ischaemia or cerebral infarction. The peptides	
Cc	are also useful as targets for drug development for the treatment of	
Cc	these diseases.	
Cc		
S0	Sequence 4 AA:	
Oy	1 HARL 4	100.0%; Score 21; DB 23; Length 4;
		Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Db	1 HARL 4	Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
RESULT 2		
ABB81511	ID ABB81511 standard; peptide; 5 AA.	
Xx		
Ac	ABB81511;	
Xx		
Dt	02-SEP-2002 (first entry)	
Xx		
De	Neural thread protein (NTP) peptide #1.	
Xx		
Kw	Neural thread protein; NTP; Harill peptide; Alzheimer's disease;	
Kw	Down's syndrome; neuroectodermal tumour; astrocytoma; glioblastoma;	
Kw	hypoxia; ischaemia; cerebral infarction.	
Xx		
Os	Homo sapiens.	
Xx		
Pn	WO200234915-A2.	
Pd	02-MAY-2002.	
Xx		
Pf	25-OCT-2001; 2001WO-US42813.	
Xx		
Pr	27-OCT-2000; 2000US-0697590.	
Xx		
Pa	(NTMO-) NTMOX PHARM CORP.	
Xx		
Pi	Fitzpatrick J, Averback P, Focht MSS, Bidiano R;	
Xx		
Dr	WPI: 2002-507998/54.	
Xx		
Pt	New Harill peptide sequences of the Neural Thread Protein, useful in	
Pt	therapeutic assays, e.g. as targets for developing drugs for treating	
Pt	Alzheimer's disease, ischemia or cerebral infarction, or for diagnosing	
Pt	these diseases -	
Xx		
PS	Claim 1; Page 29; 53pp; English.	
Xx		
Cc	The present invention describes a neural thread protein (NTP) peptide	
Cc	having an amino acid sequence selected from ABB81511 to ABB81529 and	
Cc	their homologues, which are referred collectively as Harill peptides (I).	
Cc	(I) have neuroprotective, nootropic, vasotropic and cerebroprotective	
Cc	activities, and can be used in peptide therapy. The Harill peptide	
Cc	sequences can be used as analogues for NTP in therapeutic or diagnostic	
Cc	assays by replacing NTP with the peptide in such an assay. The Harill	
Cc	peptides are also useful as a trap material in a diagnostic or	
Cc	therapeutic assay. Therefore, the Harill peptides are useful in binding	
Cc	assays, protein and antibody purification, therapeutics or diagnostics.	
Cc	In particular, the peptides are also useful for diagnosing Alzheimer's	
Cc	disease, Down's syndrome, neuroectodermal tumours, astrocytomas,	
Cc	glioblastomas, hypoxia, ischaemia or cerebral infarction. The peptides	
Cc	are also useful as targets for drug development for the treatment of	
Cc	these diseases.	
Cc		

	CC	sequences can be used as analogues for NTP in therapeutic or diagnostic assays by replacing NTP with the peptide in such an assay.
	CC	The Harill peptides are also useful as a trap material in a diagnostic or therapeutic assay.
	CC	Therefore, the Harill peptides are useful in binding assays, protein and antibody purification, therapeutics or diagnostics.
	CC	In particular, the peptides are also useful for diagnosing Alzheimer's disease, Down's syndrome, neuroectodermal tumours, astrocytomas,
	CC	glioblastomas, hypoxia, ischaemia or cerebral infarction. The peptides are also useful as targets for drug development for the treatment of these diseases.
SQ	XX	Sequence 5 AA:
Query Match		100.0%; Score 21; DB 23; Length 5;
Best Local Similarity		100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative		0; Mismatches 0; Indels 0; Gaps 0;
OY		1 HARL 4
DB		2 HARL 5
RESULT 3		
ID ABB81513		ABB81513 standard; peptide: 5 AA.
AC ABB81513;		
DJ 02-SEP-2002 (first entry)		
DE Neural thread protein (NTP) peptide #3.		
KW Neural thread protein; NTP; Harill peptide; Alzheimer's disease; Down's syndrome; neuroectodermal tumour; astrocytoma; glioblastoma; hypoxia; ischaemia; cerebral infarction.		
KM Homo sapiens.		
OS WO200234915-A2.		
PV 02-MAY-2002.		
PD 25-OCT-2001; 2001WO-US42813.		
PF 27-OCT-2000; 2000US-0697590. (NYMO-) NYMOX PHARM CORP.		
PR Fitzpatrick J, Averback P, Focht MSS, Bldiano R; WIPI; 2002-507998/54.		
DR New Harill peptide sequences of the Neural Thread Protein, useful in therapeutic assays, e.g. as targets for developing drugs for treating Alzheimer's disease, ischemia or cerebral infarction, or for diagnosing these diseases - Claim 1; Page 29; 53pp; English.		
PS The present invention describes a neural thread protein (NTP) peptide having an amino acid sequence selected from ABB81511 to ABB81519 and their homologues, which are referred collectively as Harill peptides (I). (I) have neuroprotective, nootropic, vasotrophic and cerebroprotective activities, and can be used in peptide therapy. The Harill peptide sequences can be used as analogues for NTP in therapeutic or diagnostic assays by replacing NMP with the peptide in such an assay. The Harill peptides are also useful as a trap material in a diagnostic or therapeutic assay. Therefore, the Harill peptides are useful in binding assays, protein and antibody purification, therapeutics or diagnostics. In particular, the peptides are also useful for diagnosing Alzheimer's disease, Down's syndrome, neuroectodermal tumors, astrocytomas, glioblastomas, hypoxia, ischaemia or cerebral infarction. The peptides are also useful as targets for drug development for the treatment of		

CC these diseases.
XX
SQ Sequence 5 AA:
Query Match 100.0%; Score 21; DB 23; Length 5;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 HARL 4
DB 1 HARL 4
RESULT 4
ABB81514
ID ABB81514 standard; peptide: 6 AA.
XX
AC ABB81514;
XX
DT 02-SEP-2002 (first entry)
XX
DE Neural thread protein (NTP) peptide #4.
XX
KM Neural thread protein; NTP; Harill peptide; Alzheimer's disease;
KM Down's syndrome; neuroectodermal tumour; astrocytoma; glioblastoma;
KM hypoxia; ischaemia; cerebral infarction.
XX
OS Homo sapiens.
XX
PN MO200234915-A2.
XX
PD 02-MAY-2002.
XX
PF 25-OCT-2001; 2001MO-US42813.
XX
PR 27-OCT-2000; 2000US-0697590.
XX
PT (NYMO-) NYMOX PHARM CORP.
XX
PI Fitzpatrick J, Averback P, Focht MSS, Bibiano R;
XX
DR WPI; 2002-507998/54.
XX
PT New Harill peptide sequences of the Neural Thread Protein, useful in
PT therapeutic assays, e.g. as targets for developing drugs for treating
PT Alzheimer's disease, ischemia or cerebral infarction, or for diagnosing
PT these diseases -
XX
PS Claim 1; Page 29; 53pp; English.
XX
CC The present invention describes a neural thread protein (NTP) peptide
CC having an amino acid sequence selected from ABB81511 to ABB81529 and
CC their homologues, which are referred collectively as Harill peptides (I).
CC (I) have neuroprotective, nootropic, vasotropic and cerebroprotective
CC activities, and can be used in peptide therapy. The Harill peptide
CC sequences can be used as analogues for NTP in therapeutic or diagnostic
CC assays by replacing NTP with the peptide in such an assay. The Harill
CC peptides are also useful as a trap material in a diagnostic or
CC therapeutic assay. Therefore, the Harill peptides are useful in binding
CC assays, protein and antibody purification, therapeutics or diagnostics.
CC In particular, the peptides are also useful for diagnosing Alzheimer's
CC disease, Down's syndrome, neuroectodermal tumours, astrocytomas,
CC glioblastomas, hypoxia, ischaemia or cerebral infarction. The peptides
CC are also useful as targets for drug development for the treatment of
CC these diseases.
XX
SQ Sequence 6 AA:
Query Match 100.0%; Score 21; DB 23; Length 6;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 HARL 4
DB 1 HARL 4

DB 1 HARL 4
RESULT 5
ABB81521
ID ABB81521 standard; peptide: 6 AA.
XX
AC ABB81521;
XX
DT 02-SEP-2002 (first entry)
XX
DE Neural thread protein (NTP) peptide #11.
XX
KM Neural thread protein; NTP; Harill peptide; Alzheimer's disease;
KM Down's syndrome; neuroectodermal tumour; astrocytoma; glioblastoma;
KM hypoxia; ischaemia; cerebral infarction.
XX
OS Homo sapiens.
XX
PN MO200234915-A2.
XX
PD 02-MAY-2002.
XX
PF 25-OCT-2001; 2001MO-US42813.
XX
PR 27-OCT-2000; 2000US-0697590.
XX
PT (NYMO-) NYMOX PHARM CORP.
XX
PI Fitzpatrick J, Averback P, Focht MSS, Bibiano R;
XX
DR WPI; 2002-507998/54.
XX
PT New Harill peptide sequences of the Neural Thread Protein, useful in
PT therapeutic assays, e.g. as targets for developing drugs for treating
PT Alzheimer's disease, ischemia or cerebral infarction, or for diagnosing
PT these diseases -
XX
PS Claim 1; Page 29; 53pp; English.
XX
CC The present invention describes a neural thread protein (NTP) peptide
CC having an amino acid sequence selected from ABB81511 to ABB81529 and
CC their homologues, which are referred collectively as Harill peptides (I).
CC (I) have neuroprotective, nootropic, vasotropic and cerebroprotective
CC activities, and can be used in peptide therapy. The Harill peptide
CC sequences can be used as analogues for NTP in therapeutic or diagnostic
CC assays by replacing NTP with the peptide in such an assay. The Harill
CC peptides are also useful as a trap material in a diagnostic or
CC therapeutic assay. Therefore, the Harill peptides are useful in binding
CC assays, protein and antibody purification, therapeutics or diagnostics.
CC In particular, the peptides are also useful for diagnosing Alzheimer's
CC disease, Down's syndrome, neuroectodermal tumours, astrocytomas,
CC glioblastomas, hypoxia, ischaemia or cerebral infarction. The peptides
CC are also useful as targets for drug development for the treatment of
CC these diseases.
XX
SQ Sequence 6 AA:
Query Match 100.0%; Score 21; DB 23; Length 6;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 HARL 4
DB 1 HARL 4
RESULT 6
ABB81528
ID ABB81528 standard; peptide: 6 AA.
XX
AC ABB81528;

XX 02-SEP-2002 (first entry)
DE Neural thread protein (NTP) peptide #18.
XX
XX Neural thread protein; NTP; Harill peptide; Alzheimer's disease;
KM Down's syndrome; neuroectodermal tumour; astrocytoma; glioblastoma;
XX hypoxia; ischaemia; cerebral infarction.
OS Homo sapiens.
XX
XX WO200234915-A2.
XX
XX 02-MAY-2002.
XX
XX 25-OCT-2001; 2001WO-US42813.
XX
XX 27-OCT-2000; 2000US-0697590.
XX
XX (NYMO-) NYMOX PHARM CORP.
XX
XX Fitzpatrick J, Averback P, Focht MSS, Bibiano R;
XX
XX WPI: 2002-507998/54.
XX
XX New Harill peptide sequences of the Neural Thread Protein, useful in
PT therapeutic assays, e.g. as targets for developing drugs for treating
PT Alzheimer's disease, ischemia or cerebral infarction, or for diagnosing
PT these diseases -
XX
XX Claim 1; Page 29; 53pp; English.
XX
XX The present invention describes a neural thread protein (NTP) peptide
CC having an amino acid sequence selected from ABB81511 to ABB81529 and
CC their homologues, which are referred collectively as Harill peptides (I).
CC (I) have neuroprotective, nootropic, vasotropic and cerebroprotective
CC activities, and can be used in peptide therapy. The Harill peptide
CC sequences can be used as analogues for NTP in therapeutic or diagnostic
CC assays by replacing NTP with the peptide in such an assay. The Harill
CC peptides are also useful as a trap material in a diagnostic or
CC therapeutic assay. Therefore, the Harill peptides are useful in binding
CC assays, protein and antibody purification, therapeutics or diagnostics.
CC In particular, the peptides are also useful for diagnosing Alzheimer's
CC disease, Down's syndrome, neuroectodermal tumours, astrocytomas,
CC glioblastomas, hypoxia, ischaemia or cerebral infarction. The peptides
CC are also useful as targets for drug development for the treatment of
CC these diseases.
XX
XX Sequence 6 AA:
SQ
OY 1 HARL 4
DB 1 HARL 4
Query Match 100.0%; Score 21; DB 23; Length 6;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OS Homo sapiens.
XX
XX WO200234915-A2.
XX
XX 02-MAY-2002.
XX
XX 25-OCT-2001; 2001WO-US42813.
XX
XX 27-OCT-2000; 2000US-0697590.
XX
XX (NYMO-) NYMOX PHARM CORP.
XX
XX Fitzpatrick J, Averback P, Focht MSS, Bibiano R;
XX
XX WPI: 2002-507998/54.
XX
XX New Harill peptide sequences of the Neural Thread Protein, useful in
PT therapeutic assays, e.g. as targets for developing drugs for treating
PT Alzheimer's disease, ischemia or cerebral infarction, or for diagnosing
PT these diseases -
XX
XX Claim 1; Page 29; 53pp; English.
XX
XX The present invention describes a neural thread protein (NTP) peptide
CC having an amino acid sequence selected from ABB81511 to ABB81529 and
CC their homologues, which are referred collectively as Harill peptides (I).
CC (I) have neuroprotective, nootropic, vasotropic and cerebroprotective
CC activities, and can be used in peptide therapy. The Harill peptide
CC sequences can be used as analogues for NTP in therapeutic or diagnostic
CC assays by replacing NTP with the peptide in such an assay. The Harill
CC peptides are also useful as a trap material in a diagnostic or
CC therapeutic assay. Therefore, the Harill peptides are useful in binding
CC assays, protein and antibody purification, therapeutics or diagnostics.
CC In particular, the peptides are also useful for diagnosing Alzheimer's
CC disease, Down's syndrome, neuroectodermal tumours, astrocytomas,
CC glioblastomas, hypoxia, ischaemia or cerebral infarction. The peptides
CC are also useful as targets for drug development for the treatment of
CC these diseases.
XX
XX Sequence 7 AA:
SQ
OY 1 HARL 4
DB 2 HARL 5
Query Match 100.0%; Score 21; DB 23; Length 7;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 8
ABB81517
ID ABB81517 standard; peptide: 7 AA.
XX
XX ABB81517;
XX
XX 02-SEP-2002 (first entry)
XX
XX Neural thread protein (NTP) peptide #7.
XX
XX Neural thread protein; NTP; Harill peptide; Alzheimer's disease;
KM Down's syndrome; neuroectodermal tumour; astrocytoma; glioblastoma;
XX hypoxia; ischaemia; cerebral infarction.
XX
XX Homo sapiens.
XX
XX WO200234915-A2.
XX
XX 02-MAY-2002.
XX
XX 25-OCT-2001; 2001WO-US42813.
XX
XX 27-OCT-2000; 2000US-0697590.

```

XX (NYMO-) NYMOX PHARM CORP.
PA Fitzpatrick J, Averbach P, Focht MSS, Bibiano R;
XX WPI; 2002-507998/54.
DR
XX New Harill peptide sequences of the Neural Thread Protein, useful in
PT therapeutic assays, e.g. as targets for developing drugs for treating
PT Alzheimer's disease, ischemia or cerebral infarction, or for diagnosing
PT these diseases -
XX
XX Claim 1; Page 29; 53pp; English.
XX
CC The present invention describes a neural thread protein (NTP) peptide
CC having an amino acid sequence selected from ABB81511 to ABB81529 and
CC their homologues, which are referred collectively as Harill peptides (1).
CC (1) have neuroprotective, nootropic, vasotropic and cerebroprotective
CC activities, and can be used in peptide therapy. The Harill peptide
CC sequences can be used as analogues for NTP in therapeutic or diagnostic
CC assays by replacing NTP with the peptide in such an assay. The Harill
CC peptides are also useful as a trap material in a diagnostic or
CC therapeutic assay. Therefore, the Harill peptides are useful in binding
CC assays, protein and antibody purification, therapeutics or diagnostics.
CC In particular, the peptides are also useful for diagnosing Alzheimer's
CC disease, Down's syndrome, neuroectodermal tumours, astrocytomas,
CC glioblastomas, hypoxia, ischaemia or cerebral infarction. The peptides
CC are also useful as targets for drug development for the treatment of
CC these diseases.
XX
XX Sequence 7 AA;
SO
Query Match 100.0%; Score 21; DB 23; Length 7;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
DB 2 HARL 5

RESULT 9
ABB81518
ID ABB81518 standard; peptide: 7 AA.
XX
AC ABB81518;
XX
DT 02-SEP-2002 (first entry)
XX
DE Neural thread protein (NTP) peptide #8.
XX
KW Neural thread protein; NTP; Harill peptide; Alzheimer's disease;
KW Down's syndrome; neuroectodermal tumour; astrocytoma; glioblastoma;
KW hypoxia; ischaemia; cerebral infarction.
XX
OS Homo sapiens.
XX
PN WO200234915-A2.
XX
PD 02-MAY-2002.
XX
PF 25-OCT-2001; 2001WO-US42813.
XX
PR 27-OCT-2000; 2000US-0697590.
XX
PA (NYMO-) NYMOX PHARM CORP.
XX
PI Fitzpatrick J, Averbach P, Focht MSS, Bibiano R;
XX
DR WPI; 2002-507998/54.
XX
PT New Harill peptide sequences of the Neural Thread Protein, useful in
PT therapeutic assays, e.g. as targets for developing drugs for treating

```

```

PT Alzheimer's disease, ischemia or cerebral infarction, or for diagnosing
PT these diseases -
XX
XX Claim 1; Page 29; 53pp; English.
XX
CC The present invention describes a neural thread protein (NTP) peptide
CC having an amino acid sequence selected from ABB81511 to ABB81529 and
CC their homologues, which are referred collectively as Harill peptides (1).
CC (1) have neuroprotective, nootropic, vasotropic and cerebroprotective
CC activities, and can be used in peptide therapy. The Harill peptide
CC sequences can be used as analogues for NTP in therapeutic or diagnostic
CC assays by replacing NTP with the peptide in such an assay. The Harill
CC peptides are also useful as a trap material in a diagnostic or
CC therapeutic assay. Therefore, the Harill peptides are useful in binding
CC assays, protein and antibody purification, therapeutics or diagnostics.
CC In particular, the peptides are also useful for diagnosing Alzheimer's
CC disease, Down's syndrome, neuroectodermal tumours, astrocytomas,
CC glioblastomas, hypoxia, ischaemia or cerebral infarction. The peptides
CC are also useful as targets for drug development for the treatment of
CC these diseases.
XX
XX Sequence 7 AA;
SO
Query Match 100.0%; Score 21; DB 23; Length 7;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
DB 2 HARL 5

RESULT 10
ABB81537
ID ABB81537 standard; peptide: 8 AA.
XX
AC ABB81537;
XX
DT 02-SEP-2002 (first entry)
XX
DE Neural thread protein (NTP) peptide #27.
XX
KW Neural thread protein; NTP; Harill peptide; Alzheimer's disease;
KW Down's syndrome; neuroectodermal tumour; astrocytoma; glioblastoma;
KW hypoxia; ischaemia; cerebral infarction.
XX
OS Homo sapiens.
XX
PN WO200234915-A2.
XX
PD 02-MAY-2002.
XX
PF 25-OCT-2001; 2001WO-US42813.
XX
PR 27-OCT-2000; 2000US-0697590.
XX
PA (NYMO-) NYMOX PHARM CORP.
XX
PI Fitzpatrick J, Averbach P, Focht MSS, Bibiano R;
XX
DR WPI; 2002-507998/54.
XX
PT New Harill peptide sequences of the Neural Thread Protein, useful in
PT therapeutic assays, e.g. as targets for developing drugs for treating
PT Alzheimer's disease, ischemia or cerebral infarction, or for diagnosing
PT these diseases -
XX
XX Claim 3; Page 30; 53pp; English.
XX
CC The present invention describes a neural thread protein (NTP) peptide
CC having an amino acid sequence selected from ABB81511 to ABB81529 and
CC their homologues, which are referred collectively as Harill peptides (1).
CC (1) have neuroprotective, nootropic, vasotropic and cerebroprotective

```

CC activities, and can be used in peptide therapy. The Harill peptide
CC sequences can be used as analogues for NTP in therapeutic or diagnostic
CC assays by replacing NTP with the peptide in such an assay. The Harill
CC peptides are also useful as a trap material in a diagnostic or
CC therapeutic assay. Therefore, the Harill peptides are useful in binding
CC assays, protein and antibody purification, therapeutics or diagnostics.
CC In particular, the peptides are also useful for diagnosing Alzheimer's
CC disease, Down's syndrome, neuroectodermal tumours, astrocytomas,
CC glioblastomas, hypoxia, ischaemia or cerebral infarction. The peptides
CC are also useful as targets for drug development for the treatment of
CC these diseases. The present sequence represents a specifically claimed
CC peptide from the present invention.

SQ Sequence 8 AA;

Query Match 100.0%; Score 21; DB 23; Length 8;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
|
|
|
|
DB 3 HARL 6

RESULT 11
AAG99428

ID AAG99428 standard; Peptide; 9 AA.

AC AAG99428;

DT 01-OCT-2001 (first entry)

DE PRAME derived HLA-A*0201 binding peptide SEQ ID NO: 70.

KM Vaccine; immune response; T cell response; epitope; proteasome;
KW cancer; infection.

OS Unidentified.

PN EP118860-A1.

PD 25-JUL-2001.

PF 21-JAN-2000; 2000EP-0200242.

PR 21-JAN-2000; 2000EP-0200242.

PA (UYLE-) RIJXSUNIV LEIDEN.

PA (SCIS-) SCI SEED CAPITAL INVESTMENTS BV.

PI Ossendorp F, Offringa R, Melief CJM, Kessler JH;

DR WPI; 2001-427158/46.

PT Selecting and/or producing a T cell epitope useful in a vaccine
PT comprises subjecting a precursor peptide or polypeptide to the action
PT of a 20S proteasome to determine the location of the C-terminus

PS Disclosure; Page 43; 102pp; English.

CC The present invention describes a method of producing T cell epitopes,
CC involving subjecting a precursor peptide to the action of a 20S
CC proteasome, in order to locate the C-terminus of said epitope. This can
CC be used in the production of vaccines, which can then be used to provoke
CC a T cell response in the treatment of diseases such as cancer and
CC infections. The present sequence is a peptide described in the
CC exemplification of the invention.

SQ Sequence 9 AA;

Query Match 100.0%; Score 21; DB 22; Length 9;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
|
|
|
|
DB 3 HARL 6

RESULT 12

AAG99478

ID AAG99478 standard; Peptide; 9 AA.

AC AAG99478;

DT 01-OCT-2001 (first entry)

DE PRAME derived HLA-A*0201 binding peptide SEQ ID NO: 120.

KM Vaccine; immune response; T cell response; epitope; proteasome;
KW cancer; infection.

OS Unidentified.

PN EP118860-A1.

PD 25-JUL-2001.

PF 21-JAN-2000; 2000EP-0200242.

PR 21-JAN-2000; 2000EP-0200242.

PA (UYLE-) RIJXSUNIV LEIDEN.

PA (SCIS-) SCI SEED CAPITAL INVESTMENTS BV.

PI Ossendorp F, Offringa R, Melief CJM, Kessler JH;

DR WPI; 2001-427158/46.

PT Selecting and/or producing a T cell epitope useful in a vaccine
PT comprises subjecting a precursor peptide or polypeptide to the action
PT of a 20S proteasome to determine the location of the C-terminus

PS Disclosure; Page 62; 102pp; English.

CC The present invention describes a method of producing T cell epitopes,
CC involving subjecting a precursor peptide to the action of a 20S
CC proteasome, in order to locate the C-terminus of said epitope. This can
CC be used in the production of vaccines, which can then be used to provoke
CC a T cell response in the treatment of diseases such as cancer and
CC infections. The present sequence is a peptide described in the
CC exemplification of the invention.

SQ Sequence 9 AA;

Query Match 100.0%; Score 21; DB 22; Length 9;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
|
|
|
|
DB 6 HARL 9

RESULT 13

AAG99452

ID AAG99452 standard; Peptide; 10 AA.

AC AAG99452;

DT 01-OCT-2001 (first entry)

DE PRAME derived HLA-A*0201 binding peptide SEQ ID NO: 94.

KM Vaccine; immune response; T cell response; epitope; proteasome;
KW cancer; infection.

XX Unidentified.
 OS
 XX
 PN EP118860-A1.
 XX
 PD 25-JUL-2001.
 XX
 PF 21-JAN-2000; 2000EP-0200242.
 XX
 PR 21-JAN-2000; 2000EP-0200242.
 XX
 PA (UYLE-) RIJKSUNIV LEIDEN.
 (SCIS-) SCI SEED CAPITAL INVESTMENTS BV.
 XX
 PI Ossendorf F, Offringa R, Melief CJM, Kessler JH;
 DR WPI; 2001-427158/46.
 XX
 PT Selecting and/or producing a T cell epitope useful in a vaccine
 PT comprises subjecting a precursor peptide or polypeptide to the action
 of a 20S proteasome to determine the location of the C-terminus
 XX
 PS Disclosure; Page 52; 102pp; English.
 CC The present invention describes a method of producing T cell epitopes,
 CC involving subjecting a precursor peptide to the action of a 20S
 CC proteasome, in order to locate the C-terminus of said epitope. This can
 CC be used in the production of vaccines, which can then be used to provoke
 CC a T cell response in the treatment of diseases such as cancer and
 CC infections. The present sequence is a peptide described in the
 CC exemplification of the invention.
 CC
 SQ Sequence 10 AA;
 XX
 Query Match 100.0%; Score 21; DB 22; Length 10;
 Best Local Similarity 100.0%; Pred. No. 74;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 HARL 4
 ||||
 DB 3 HARL 6
 RESULT 14
 ABB81536
 ID ABB81536 standard; peptide; 10 AA.
 XX
 AC ABB81536;
 XX
 DT 02-SEP-2002 (first entry)
 XX
 DE Neural thread protein (NTP) peptide #26.
 XX
 KW Neural thread protein; NTP; Harill peptide; Alzheimer's disease;
 KM Down's syndrome; neuroectodermal tumour; astrocytoma; glioblastoma;
 KM hypoxia; ischaemia; cerebral infarction.
 XX
 OS Homo sapiens.
 XX
 PN WO200234915-A2.
 XX
 PD 02-MAY-2002.
 XX
 PF 25-OCT-2001; 2001WO-US42813.
 XX
 PR 27-OCT-2000; 2000US-0697590.
 XX
 PA (NYMO-) NYMOX PHARM CORP.
 XX
 PI Fitzpatrick J, Averbach P, Focht MSS, Bibiano R;
 DR WPI; 2002-507998/54.
 XX

PT New Harill peptide sequences of the Neural Thread Protein, useful in
 PT therapeutic assays, e.g. as targets for developing drugs for treating
 PT Alzheimer's disease, ischemia or cerebral infarction, or for diagnosing
 PT these diseases
 XX
 PS Claim 3; Page 30; 53pp; English.
 XX
 CC The present invention describes a neural thread protein (NTP) peptide
 CC having an amino acid sequence selected from ABB81511 to ABB81529 and
 CC their homologues, which are referred collectively as Harill peptides (1).
 CC (1) have neuroprotective, neurotropic, vasotropic and cerebroprotective
 CC activities, and can be used in peptide therapy. The Harill peptide
 CC sequences can be used as analogues for NTP in therapeutic or diagnostic
 CC assays by replacing NTP with the peptide in such an assay. The Harill
 CC peptides are also useful as a trap material in a diagnostic or
 CC therapeutic assay. Therefore, the Harill peptides are useful in binding
 CC assays, protein and antibody purification, therapeutics or diagnostics.
 CC In particular, the peptides are also useful for diagnosing Alzheimer's
 CC disease, Down's syndrome, neuroectodermal tumours, astrocytomas,
 CC glioblastomas, hypoxia, ischaemia or cerebral infarction. The peptides
 CC are also useful as targets for drug development for the treatment of
 CC these diseases. The present sequence represents a specifically claimed
 CC peptide from the present invention.
 CC
 SQ Sequence 10 AA;
 XX
 Query Match 100.0%; Score 21; DB 23; Length 10;
 Best Local Similarity 100.0%; Pred. No. 74;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 HARL 4
 ||||
 DB 5 HARL 8
 RESULT 15
 AAB31270
 ID AAB31270 standard; peptide; 12 AA.
 XX
 AC AAB31270;
 XX
 DT 20-APR-2001 (first entry)
 XX
 DE Peptide ligand which binds to human MUC1 protein.
 XX
 KW MUC1; epithelial cell mucin; tumour; MUC1 ligand; cancer.
 XX
 OS Synthetic.
 XX
 PN WO200077031-A2.
 XX
 PD 21-DEC-2000.
 XX
 PF 15-JUN-2000; 2000WO-CA00711.
 XX
 PR 15-JUN-1999; 99US-0139263.
 XX
 PA (UYHE-) UNIV HEALTH NETWORK.
 XX
 PI Garlepy J, Yang S;
 DR WPI; 2001-091388/10.
 XX
 PT Novel ligands that binds to MUC1, an epithelial cell mucin, useful for
 PT diagnosing, monitoring, treating and preventing cancer
 XX
 PS Claim 15; Page 42; 60pp; English.
 XX
 CC The present sequence represents a peptide ligand which binds to human
 CC MUC1 protein. MUC1 is an epithelial cell mucin, which is found on, and
 CC shed from, the surface of many tumours. MUC1 peptide ligands were
 CC isolated using a phage display technique using MUC1 tandem repeats as
 CC the target. The MUC1 tandem target repeats comprise repeats of the

CC peptide ABB1258. The ligands are useful for detecting the presence of
CC MUC1 in a sample, to treat or prevent cancer associated with MUC1 and
CC to prepare a medicament or diagnostic agent to treat, prevent or detect
CC cancer associated with MUC1. By assaying for the binding between the
CC MUC1 ligand and MUC1 in a sample diagnosis or monitoring of cancer can
CC be carried out. The ligands may also be used to prepare antibodies.

XX
SQ Sequence 12 AA;

Query Match 100.0%; Score 21; DB 22; Length 12;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
|||||
DB 7 HARL 10

RESULT 16

ABB81533
ID ABB81533 standard; peptide; 12 AA.

XX
AC ABB81533;

XX
DT 02-SEP-2002 (first entry)

XX
DE Neural thread protein (NTP) peptide #23.

XX
KW Neural thread protein; NTP; Har11 peptide; Alzheimer's disease;

KW Down's syndrome; neuroectodermal tumour; astrocytoma; glioblastoma;

KW hypoxia; ischemia; cerebral infarction.

XX
OS Homo sapiens.

XX
PN WO200234915-A2.

XX
PD 02-MAY-2002.

XX
PE 25-OCT-2001; 2001WO-US42813.

XX
PR 27-OCT-2000; 2000US-0697590.

XX
PA (NYMO-) NYMOX PHARM CORP.

XX
PI Fitzpatrick J, Averback P, Focht MSS, Bibiano R;

XX
DR WPI; 2002-507998/54.

XX
PT New Har11 peptide sequences of the Neural Thread Protein, useful in
PT therapeutic assays, e.g. as targets for developing drugs for treating
PT Alzheimer's disease, ischemia or cerebral infarction, or for diagnosing
PT these diseases -

XX
PS Claim 3; Page 30; 53pp; English.

CC The present invention describes a neural thread protein (NTP) peptide
CC having an amino acid sequence selected from ABB81511 to ABB81529 and
CC their homologues, which are referred collectively as Har11 peptides (I).
CC (I) have neuroprotective, nootropic, vasotropic and cerebroprotective
CC activities, and can be used in peptide therapy. The Har11 peptide
CC sequences can be used as analogues for NTP in therapeutic or diagnostic
CC assays by replacing NTP with the peptide in such an assay. The Har11
CC peptides are also useful as a trap material in a diagnostic or
CC therapeutic assay. Therefore, the Har11 peptides are useful in binding
CC assays, protein and antibody purification, therapeutics or diagnostics.
CC In particular, the peptides are also useful for diagnosing Alzheimer's
CC disease, Down's syndrome, neuroectodermal tumours, astrocytomas,
CC glioblastomas, hypoxia, ischemia or cerebral infarction. The peptides
CC are also useful as targets for drug development for the treatment of
CC these diseases. The present sequence represents a specifically claimed
CC peptide from the present invention.

XX
SQ Sequence 12 AA;

Query Match 100.0%; Score 21; DB 23; Length 12;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
|||||
DB 2 HARL 5

RESULT 17

ABB81532
ID ABB81532 standard; peptide; 13 AA.

XX
AC ABB81532;

XX
DT 02-SEP-2002 (first entry)

XX
DE Neural thread protein (NTP) peptide #22.

XX
KW Neural thread protein; NTP; Har11 peptide; Alzheimer's disease;

KW Down's syndrome; neuroectodermal tumour; astrocytoma; glioblastoma;

KW hypoxia; ischemia; cerebral infarction.

XX
OS Homo sapiens.

XX
PN WO200234915-A2.

XX
PD 02-MAY-2002.

XX
PE 25-OCT-2001; 2001WO-US42813.

XX
PR 27-OCT-2000; 2000US-0697590.

XX
PA (NYMO-) NYMOX PHARM CORP.

XX
PI Fitzpatrick J, Averback P, Focht MSS, Bibiano R;

XX
DR WPI; 2002-507998/54.

XX
PT New Har11 peptide sequences of the Neural Thread Protein, useful in
PT therapeutic assays, e.g. as targets for developing drugs for treating
PT Alzheimer's disease, ischemia or cerebral infarction, or for diagnosing
PT these diseases -

XX
PS Claim 3; Page 29; 53pp; English.

CC The present invention describes a neural thread protein (NTP) peptide
CC having an amino acid sequence selected from ABB81511 to ABB81529 and
CC their homologues, which are referred collectively as Har11 peptides (I).
CC (I) have neuroprotective, nootropic, vasotropic and cerebroprotective
CC activities, and can be used in peptide therapy. The Har11 peptide
CC sequences can be used as analogues for NTP in therapeutic or diagnostic
CC assays by replacing NTP with the peptide in such an assay. The Har11
CC peptides are also useful as a trap material in a diagnostic or
CC therapeutic assay. Therefore, the Har11 peptides are useful in binding
CC assays, protein and antibody purification, therapeutics or diagnostics.
CC In particular, the peptides are also useful for diagnosing Alzheimer's
CC disease, Down's syndrome, neuroectodermal tumours, astrocytomas,
CC glioblastomas, hypoxia, ischemia or cerebral infarction. The peptides
CC are also useful as targets for drug development for the treatment of
CC these diseases. The present sequence represents a specifically claimed
CC peptide from the present invention.

XX
SQ Sequence 13 AA;

Query Match 100.0%; Score 21; DB 23; Length 13;
Best Local Similarity 100.0%; Pred. No. 96;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
|||||
DB 9 HARL 12

RESULT 18
 AAM00412
 ID AAM00412 standard; Peptide: 14 AA.
 XX
 AC AAM00412;
 XX
 DT 01-OCT-2001 (first entry)
 XX
 DE Human protein fragment SEQ ID NO: 960.
 XX
 KM Human; single nucleotide polymorphism: SNP; paternity test;
 KM forensic test; aberrant protein expression.
 XX
 OS Homo sapiens.
 XX
 PN WO200151670-A2.
 XX
 PD 19-JUL-2001.
 XX
 PE 05-JAN-2001; 2001WO-US00322.
 XX
 PR 07-JAN-2000; 2000US-0174962.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Shinkets RA, Leach MD;
 XX
 DR WPI: 2001-451871/48.
 DR N-PSDB; AAH89529.
 XX
 PT Isolated human polynucleotides containing single nucleotide
 PT polymorphisms, useful for the treatment and diagnosis of e.g. cancer,
 PT infection and diabetes -
 XX
 PS Disclosure: Page 379; 475pp; English.
 XX
 CC The present invention relates to human nucleic acids containing single
 CC nucleotide polymorphisms (SNPs). These can be used in forensic and
 CC paternity tests, and to aid in the treatment of diseases associated with
 CC aberrant protein expression, including cancer, amyloidosis, diabetes,
 CC Alzheimer's disease, Down's syndrome, oedema, lupus (SLE), vasculitis,
 CC glomerulonephritis, haemolytic anaemia, thrombocytopenia, arthritis,
 CC meningitis, muscular disorders, dementia, neurological diseases, tubercous
 CC sclerosis, male infertility, hypercalcaemia, blood pressure disorders,
 CC osteoporosis, pathogenic infections, hypercholesterolaemia, obesity or
 CC autoimmunity. The present sequence is a peptide encoded by a
 CC polymorphism-containing oligonucleotide fragment of the invention.
 XX
 SQ Sequence 14 AA;
 XX
 QY 1 HARL 4
 IIII
 Db 5 HARL 8
 XX
 Query Match 100.0%; Score 21; DB 22; Length 14;
 Best Local Similarity 100.0%; Pred. NO. 1e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX
 OS Homo sapiens.
 XX
 PN WO200151670-A2.
 XX
 PD 19-JUL-2001.
 XX
 PE 05-JAN-2001; 2001WO-US00322.
 XX
 PR 07-JAN-2000; 2000US-0174962.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Shinkets RA, Leach MD;
 XX
 DR WPI: 2001-451871/48.
 DR N-PSDB; AAH89561.
 XX
 PT Isolated human polynucleotides containing single nucleotide
 PT polymorphisms, useful for the treatment and diagnosis of e.g. cancer,
 PT infection and diabetes -
 XX
 PS Disclosure: Page 388; 475pp; English.
 XX
 CC The present invention relates to human nucleic acids containing single
 CC nucleotide polymorphisms (SNPs). These can be used in forensic and
 CC paternity tests, and to aid in the treatment of diseases associated with
 CC aberrant protein expression, including cancer, amyloidosis, diabetes,
 CC Alzheimer's disease, Down's syndrome, oedema, lupus (SLE), vasculitis,
 CC glomerulonephritis, haemolytic anaemia, thrombocytopenia, arthritis,
 CC meningitis, muscular disorders, dementia, neurological diseases, tubercous
 CC sclerosis, male infertility, hypercalcaemia, blood pressure disorders,
 CC osteoporosis, pathogenic infections, hypercholesterolaemia, obesity or
 CC autoimmunity. The present sequence is a peptide encoded by a
 CC polymorphism-containing oligonucleotide fragment of the invention.
 XX
 SQ Sequence 14 AA;
 XX
 QY 1 HARL 4
 IIII
 Db 5 HARL 8
 XX
 Query Match 100.0%; Score 21; DB 22; Length 14;
 Best Local Similarity 100.0%; Pred. NO. 1e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 20
 ABB81534
 ID ABB81534 standard; peptide: 14 AA.
 XX
 AC ABB81534;
 XX
 DT 02-SEP-2002 (first entry)
 XX
 DE Neural thread protein (NTP) peptide #24.
 XX
 KM Neural thread protein; NTP; Har11 peptide; Alzheimer's disease;
 KM Down's syndrome; neuroectodermal tumour; astrocytoma; glioblastoma;
 KM hypoxia; ischaemia; cerebral infarction.
 XX
 OS Homo sapiens.
 XX
 PN WO200234915-A2.
 XX
 PD 02-MAY-2002.
 XX
 PE 25-OCT-2001; 2001WO-US42813.
 XX
 PR 27-OCT-2000; 2000US-0697590.
 XX
 PA (NYMO-) NYMOX PHARM CORP.
 XX

PI Fitzpatrick J, Averback P, Focht MSS, Bibiano R;
XX MPI; 2002-507998/54.
XX
XX
PT New Harill peptide sequences of the Neural Thread Protein, useful in
PT therapeutic assays, e.g. as targets for developing drugs for treating
PT Alzheimer's disease, ischemia or cerebral infarction, or for diagnosing
PT these diseases -
XX
XX
PS Claim 3; Page 30; 53pp; English.
XX
XX
CC The present invention describes a neural thread protein (NTP) peptide
CC having an amino acid sequence selected from ABB81511 to ABB81529 and
CC their homologues, which are referred collectively as Harill peptides (1).
CC (1) have neuroprotective, nootropic, vasotropic and cerebroprotective
CC activities, and can be used in peptide therapy. The Harill peptide
CC sequences can be used as analogues for NTP in therapeutic or diagnostic
CC assays by replacing NTP with the peptide in such an assay. The Harill
CC peptides are also useful as a trap material in a diagnostic or
CC therapeutic assay. Therefore, the Harill peptides are useful in binding
CC assays, protein and antibody purification, therapeutics or diagnostics.
CC In particular, the peptides are also useful for diagnosing Alzheimer's
CC disease, Down's syndrome, neuroectodermal tumours, astrocytomas,
CC glioblastomas, hypoxia, ischaemia or cerebral infarction. The peptides
CC are also useful as targets for drug development for the treatment of
CC these diseases. The present sequence represents a specifically claimed
CC peptide from the present invention.
XX
SQ Sequence 14 AA;
XX
Query Match 100.0%; Score 21; DB 23; Length 14;
Best Local Similarity 100.0%; Pred. No. 1e+02; Indels 0; Gaps 0;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 HARL 4
DB 5 HARL 8
XX
RESULT 21
AAV79924
ID AAV7924 standard; Peptide; 15 AA.
XX
XX
AC AAV79924;
XX
DT 10-MAY-2000 (first entry)
XX
DE Human papillomavirus 16 E2 derived peptide M36A.
XX
KM Papillomavirus; PV; infection; cell proliferation; E2; peptidomimetic;
KM E1; antiviral; virucide; cytostatic; antiproliferative; dermatological;
KM preneoplastic lesion; neoplastic lesion; cutaneous lesion; wart;
KM epidermodysplasia verruciformis; anorectal carcinoma.
XX
OS Human papillomavirus.
OS Synthetic.
XX
XX
PN WO200001720-A2.
XX
PD 13-JAN-2000.
XX
XX
PF 02-JUL-1999; 99WO-US15144.
XX
XX
PR 02-JUL-1998; 98US-0091661.
XX
XX
PA (HARD) HARVARD COLLEGE.
XX
PI Howley P, Benson J, Kasukawa H;
XX
XX
DR MPI; 2000-171001/15.
XX
PT Use of papillomavirus E2 protein peptidomimetics for treating
PT papillomavirus-infected cells and papillomavirus-induced conditions in

PT mammals by inhibiting E1-E2 interaction -
XX
XX
PS Disclosure; Fig 10; 110pp; English.
XX
XX
CC The present invention describes the use of a small organic compound (A)
CC which competitively inhibits interaction of a papillomavirus (PV) E2
CC protein with a PV E1 protein for treating a cell infected with PV or a
CC mammal with a PV-induced condition. (A) has antiviral, virucide,
CC cytostatic, antiproliferative and dermatological activities. Methods
CC from the present invention can be used to treat PV-induced conditions
CC including growth of PV preneoplastic and neoplastic lesions, cutaneous
CC lesions chosen from warts and other benign cutaneous lesions, plantar
CC warts (verruca plantaris), common warts (verruca plana), Butcher's
CC common warts, flat warts, genital warts (condyloma acuminatum) and
CC epidermodysplasia verruciformis, laryngeal, oral, pharyngeal,
CC oesophageal and other upper airway papilloma or vaginal, cervical,
CC vulvar, penile and anorectal carcinoma. The E2 inhibitors may also be
CC used to treat epithelial and internal fibropapillomas in animals.
CC The present sequence represents a peptide sequence used in the
CC exemplification of the present invention.
XX
SQ Sequence 15 AA;
XX
Query Match 100.0%; Score 21; DB 21; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 HARL 4
DB 3 HARL 6
XX
RESULT 22
ABB81535
ID ABB81535 standard; peptide; 15 AA.
XX
XX
AC ABB81535;
XX
DT 02-SEP-2002 (first entry)
XX
DE Neural thread protein (NTP) peptide #25.
XX
KM Neural thread protein; NTP; Harill peptide; Alzheimer's disease;
KM Down's syndrome; neuroectodermal tumour; astrocytoma; glioblastoma;
KM hypoxia; ischaemia; cerebral infarction.
XX
OS Homo sapiens.
XX
PN WO200234915-A2.
XX
XX
PD 02-MAY-2002.
XX
PF 25-OCT-2001; 2001WO-US42813.
XX
XX
PR 27-OCT-2000; 2000US-0697590.
XX
XX
PA (NYMO-) NYMOX PHARM CORP.
XX
PI Fitzpatrick J, Averback P, Focht MSS, Bibiano R;
XX
XX
DR MPI; 2002-507998/54.
XX
XX
XX
PT New Harill peptide sequences of the Neural Thread Protein, useful in
PT therapeutic assays, e.g. as targets for developing drugs for treating
PT Alzheimer's disease, ischemia or cerebral infarction, or for diagnosing
PT these diseases -
XX
XX
PS Claim 3; Page 30; 53pp; English.
XX
XX
CC The present invention describes a neural thread protein (NTP) peptide
CC having an amino acid sequence selected from ABB81511 to ABB81529 and
CC their homologues, which are referred collectively as Harill peptides (1).
CC (1) have neuroprotective, nootropic, vasotropic and cerebroprotective

CC activities, and can be used in peptide therapy. The Harill peptide
CC sequences can be used as analogues for NTP in therapeutic or diagnostic
CC assays by replacing NTP with the peptide in such an assay. The Harill
CC peptides are also useful as a trap material in a diagnostic or
CC therapeutic assay. Therefore, the Harill peptides are useful in binding
CC assays, protein and antibody purification, therapeutics or diagnostics.
CC In particular, the peptides are also useful for diagnosing Alzheimer's
CC disease, Down's syndrome, neuroectodermal tumours, astrocytomas,
CC glioblastomas, hypoxia, ischaemia or cerebral infarction. The peptides
CC are also useful as targets for drug development for the treatment of
CC these diseases. The present sequence represents a specifically claimed
CC peptide from the present invention.

XX
SQ Sequence 15 AA;

Query Match 100.0%; Score 21; DB 23; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
||||
DB 11 HARL 14

RESULT 23

ABB81530
ID ABB81530 standard; peptide; 16 AA.

AC ABB81530;

XX 02-SEP-2002 (first entry)

DE Neural thread protein (NTP) peptide #20.

XX Neural thread protein; NTP; Harill peptide; Alzheimer's disease;

KM Down's syndrome; neuroectodermal tumour; astrocytoma; glioblastoma;

XX hypoxia; ischaemia; cerebral infarction.

OS Homo sapiens.

PN W0200234915-A2.

XX 02-MAY-2002.

PF 25-OCT-2001; 2001MO-US42813.

XX 27-OCT-2000; 2000US-0697590.

PA (NYMO-) NYMOX PHARM CORP.

XX Fitzpatrick J, Averbach P, Focht MSS, Biblano R;

DR WPI; 2002-507998/54.

XX

PT New Harill peptide sequences of the Neural Thread Protein, useful in

PT therapeutic assays, e.g. as targets for developing drugs for treating

PT Alzheimer's disease, ischemia or cerebral infarction, or for diagnosing

PT these diseases -

XX

PS Claim 3; Page 29; 53pp; English.

XX

CC The present invention describes a neural thread protein (NTP) peptide

CC having an amino acid sequence selected from ABB81511 to ABB81529 and

CC their homologues, which are referred collectively as Harill peptides (I).

CC glioblastomas, hypoxia, ischaemia or cerebral infarction. The peptides
CC are also useful as targets for drug development for the treatment of
CC these diseases. The present sequence represents a specifically claimed
CC peptide from the present invention.

XX
SQ Sequence 16 AA;

Query Match 100.0%; Score 21; DB 23; Length 16;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
||||
DB 2 HARL 5

RESULT 24

AAR80293
ID AAR80293 standard; Protein; 17 AA.

AC AAR80293;

XX 19-JAN-1996 (first entry)

DE Trehalose releasing enzyme internal fragment.

XX Enzyme; trehalose; alpha-glucosyltrehalose; alpha-maltosyltrehalose;

KW alpha-maltotriosyltrehalose; alpha-maltotetraosyltrehalose;

KW maltopentaosyltrehalose; sweetener; taste-improving agent;

KW stabiliser; filler; excipient; adjuvant; foodstuffs; cosmetics;

XX pharmaceuticals.

OS Rhizobium sp. M11.

PN EP671470-A2.

XX 13-SEP-1995.

PF 07-MAR-1995; 95EP-0301474.

XX 07-MAR-1994; 94JP-0059840.

PR 07-MAR-1994; 94JP-0059834.

XX (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.

PA Hattori K, Kubota M, Sugimoto T, Tsusaki K;

PI WPI; 1995-312772/41.

XX

DR DNA encoding a trehalose releasing enzyme - which releases trehalose

PT from a non-reducing saccharide having a trehalose structure as an

PT end unit.

XX

PS Example 2; Page 26; 45pp; English.

XX

CC The trehalose releasing enzyme can be used for the preparation of

CC trehalose with high yields and efficiency from non-reducing

CC saccharides such as alpha-glucosyltrehalose, alpha-maltosyltrehalose,

CC alpha-maltotriosyltrehalose, alpha-maltotetraosyltrehalose and

CC maltopentaosyltrehalose. The trehalose can be used as a sweetener,

CC taste-improving agent, quality-improving agent, stabiliser, filler,

CC excipient or adjuvant in food products cosmetics and pharmaceuticals.

CC This is an internal fragment of the enzyme.

XX

SQ Sequence 17 AA;

Query Match 100.0%; Score 21; DB 16; Length 17;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
||||
DB 11 HARL 14

RESULT 25
 AAY58899
 ID AAY58899 standard; Peptide; 25 AA.
 XX
 AC AAY58899;
 XX
 DT 23-MAY-2000 (first entry)
 XX
 DE Polyketide polylinker-encoded peptide Ave02.
 XX
 KM Polyketide synthase; avermectin; insecticide; mutant; mutein.
 XX
 OS Streptomyces avermitilis.
 OS Synthetic.
 XX
 PN WO200001827-A2.
 XX
 PD 13-JAN-2000.
 XX
 PF 06-JUL-1999; 99WO-GB02158.
 XX
 PR 06-JUL-1998; 98GB-0014622.
 XX
 PA (BIOT-) BIOTICA TECHNOLOGY LTD.
 PA (PFIZ) PFIZER INC.
 XX
 PI Kellenberger JL, Leadlay PF, Staunton J, Stutzman-Engwall KJ;
 PI McArthur HAI;
 DR WPI: 2000-182117/16.
 XX
 PT Mutated Type I polyketide synthase containing a polylinker site in an
 PT extension module for replacement of a reductive loop sequence, for
 PT producing polyketides, e.g. B1 avermectin -
 XX
 PS Disclosure: Fig 1; 75pp; English.
 XX
 CC The present sequence is that of the peptide region of a Type I
 CC polyketide synthase (PKS) that is encoded by a DNA polylinker
 CC incorporating NsiI, Bsu36I, NheI and HpaI restriction enzyme
 CC sites. The polylinker is inserted into the reductive loop of
 CC module 2 of the avermectin PKS gene. The invention relates to
 CC nucleic acid molecules encoding at least part of a Type I PKS.
 CC These have a polylinker with multiple restriction enzyme sites
 CC in place of 1 or more PKS genes encoding enzymes associated with
 CC reduction. Also provided are PKS in which the reductive loop in a
 CC selected module of the Type I PKS is replaced with the equivalent
 CC segment from the same or different PKS gene cluster or by a mutated
 CC or synthetic segment. Vectors and host cells, and methods for
 CC producing novel polyketides by culturing host cells are claimed.
 CC The polyketides obtained are useful as antibiotics and insecticides.
 CC
 XX
 SO Sequence 25 AA;
 Query Match 100.0%; Score 21; DB 21; Length 25;
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 HARL 4
 DB 8 HARL 11
 RESULT 26
 AAM94846
 ID AAM94846 standard; Protein; 30 AA.
 XX
 AC AAM94846;
 XX
 DT 07-MAY-1999 (first entry)
 XX

DE Bait region of wild-type human alpha-2-macroglobulin (h-alpha-2M).
 XX
 KM Protein degrading enzyme; inhibitor; modified; bait region; enzyme;
 KM alpha-2-macroglobulin; alpha-1-macroglobulin; pregnancy zone protein;
 KM PzP; medicine; human; h-alpha-2M.
 XX
 OS Homo sapiens.
 OS
 PN JP11032777-A.
 XX
 PD 09-FEB-1999.
 XX
 PF 24-JUL-1997; 97JP-0212701.
 XX
 PR 24-JUL-1997; 97JP-0212701.
 XX
 PA (MEIP) MEIJI MILK PROD CO LTD.
 XX
 DR WPI: 1999-183834/16.
 DR N-PSDB; AAX05799.
 XX
 PT New protein degrading enzyme inhibitor and its preparation - causes
 PT no rejection or immunological abnormal response and is useful in
 PT medicine
 XX
 PS Examples: Fig 1A; 7pp; Japanese.
 XX
 CC The invention relates to an inhibitor of protein degrading enzyme. The
 CC inhibitor is partially modified at a bait region of the enzyme. In which
 CC the peptide sequence of the bait region is derived from alpha-2-
 CC macroglobulin, alpha-1-macroglobulin or pregnancy zone protein (PzP). The
 CC inhibitor is useful in medicine. Administration of the inhibitor to an
 CC animal causes no rejection or immunological abnormal response with short
 CC period of half life in the living body. The present sequence represents a
 CC the wild-type bait region of human alpha-2-macroglobulin (h-alpha-2M).
 CC
 XX
 SO Sequence 30 AA;
 Query Match 100.0%; Score 21; DB 20; Length 30;
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 HARL 4
 DB 20 HARL 23
 RESULT 27
 AAY56381
 ID AAY56381 standard; protein; 30 AA.
 XX
 AC AAY56381;
 XX
 DT 14-FEB-2000 (first entry)
 XX
 DE Neisseria gonorrhoeae yJeq family member motif 4 peptide sequence.
 XX
 KM yJeq family; bacterial; growth; computer based algorithm; antibiotic;
 KM infection; identification.
 XX
 OS Neisseria gonorrhoeae.
 OS
 PN WO954473-A2.
 XX
 PD 28-OCT-1999.
 XX
 PF 20-APR-1999; 99WO-EP02636.
 XX
 PR 22-APR-1998; 98GB-0008366.
 XX
 PA (GLAX) GLAXO GROUP LTD.
 XX
 PI Arigoni F, Edgerton MD, Loferer H, Peitsch MC;

CC	parent polypeptide to impede, abrogate or otherwise alter at least one
CC	function associated with the parent polypeptide and for inducing an
CC	immune response against a pathogen or cancer. Also included are a
CC	synthetic polynucleotide encoding and a computer system for
CC	designing the synthetic polypeptides. The synthetic polypeptides and
CC	polynucleotides are referred to as a Savine. The synthetic polypeptide is
CC	useful for modulating immune responses preferably directed against a
CC	pathogen or a cancer, (e.g., cancers of the lung, breast, ovary, cervix,
CC	colon, head and neck, pancreas, prostate, stomach, bladder, kidney, bone
CC	liver, oesophagus, brain, testicle, uterus), as potentiating agents.
CC	Compositions comprising the polypeptide may be used in the treatment or
CC	prophylaxis against viral (such as infections caused by HIV (human
CC	immunodeficiency virus), hepatitis, influenza, Japanese encephalitis
CC	virus, Epstein-Barr virus and respiratory syncytial virus), bacterial
CC	(e.g., infections caused by Neisseria, Meningococcal, Haemophilus,
CC	Salmonella, Streptococcal, Legionella and Mycobacterium) or parasitic
CC	(e.g., infections caused by Plasmodium, Schistosoma, Leishmania,
CC	Trypanosoma, Toxoplasma and Giardia) infections. The present
CC	sequence is a peptide derived from a parent protein used to
CC	construct a savine of the invention.
XX	
SO	Sequence 30 AA:
Qy	1 HARL 4
Db	16 HARL 19
RESULT 29	
AAU85096	
AAU85096	standard; Peptide: 30 AA.
XX	
AC	AAU85096;
XX	
DT	08-MAY-2002 (first entry)
XX	
DE	Human PRAME segment 32.
XX	
XX	Savine; vaccine; cancer; viral infection; HIV; hepatitis C virus;
KW	viral infection; human immunodeficiency virus; melanoma;
KW	bacterial infection; Salmonella; Legionella; parasitic infection;
KW	Trypanosoma; Toxoplasma; Giardia.
XX	
OS	Homo sapiens.
XX	
PN	WO200190197-A1.
XX	
PD	29-NOV-2001.
XX	
PF	25-MAY-2001; 2001WO-AU00622.
XX	
PR	26-MAY-2000; 2000AU-0007761.
XX	
PA	(AUSU) UNIV AUSTRALIAN NAT.
XX	
PI	Thomson SA, Ramsdew IA;
XX	
DR	WPI: 2002-147575/19.
DR	N-PSDB: ABK36916.
PT	New synthetic polypeptides having several different segments of at
PT	least one parent polypeptide linked together differently compared to
PT	the linkage in the parent polypeptide, for inducing immune response
PT	against a pathogen or cancer
XX	
PS	Example 3; Fig 27; 364pp; English.
XX	
CC	The invention relates to a new synthetic polypeptide (I) comprising
CC	several different segments of at least one parent polypeptide linked

CC together in a different relationship relative to their linkage in the
CC parent polypeptide to impede, abrogate or otherwise alter at least one
CC function associated with the parent polypeptide and for inducing an
CC immune response against a pathogen or cancer. Also included are a
CC synthetic polynucleotide encoding and a computer system for
CC designing the synthetic polypeptides. The synthetic polypeptides and
CC polynucleotides are referred to as a Savine. The synthetic polypeptide is
CC useful for modulating immune responses preferably directed against a
CC pathogen or a cancer, (e.g., cancers of the lung, breast, ovary, cervix,
CC colon, head and neck, pancreas, prostate, stomach, bladder, kidney, bone
CC liver, oesophagus, brain, testicle, uterus), as potentiating agents.
CC Compositions comprising the polypeptide may be used in the treatment or
CC prophylaxis against viral (such as infections caused by HIV (human
CC immunodeficiency virus), hepatitis, influenza, Japanese encephalitis
CC virus, Epstein-Barr virus and respiratory syncytial virus), bacterial
CC (e.g., infections caused by Neisseria, Meningococcal, Hemophilus,
CC Salmonella, Streptococcal, Legionella and Mycobacterium) or parasitic
CC (e.g., infections caused by Plasmodium, Schistosoma, Leishmania,
CC Trypanosoma, Toxoplasma and Giardia) infections. The present
CC sequence is a peptide derived from a parent protein used to
CC construct a savine of the invention.

CC Sequence 30 AA;

Query Match 100.0%; Score 21; DB 23; Length 30;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
1111
DB 1 HARL 4

RESULT 30
AAM94447

ID AAM94447 standard; peptide; 31 AA.

AC AAM94447;

DT 15-APR-1999 (first entry)

DE Mutant propeptidic linker region for MMP-11, PAP-284.

KW R1c1n-like toxin; cancer; viral infection; parasitic infection;
KW linker; B chain; A chain; protease; fungal infection; malaria;
KW leucocyte proliferation; cytomegalovirus; herpes; hepatitis;
KW rhinovirus; laryngeotracheitis; polioyelitis; varicella zoster;
KW cystic fibrosis; multiple sclerosis.

OS Unidentified.
Synthetic.

PN MO9849311-A2.

XX 05-NOV-1998.

PD 30-APR-1998; 98WO-CA00394.

PF 29-OCT-1997; 97US-0063715.

PR 30-APR-1997; 97US-0045148.

XX (DNOV-) DE NOVO ENZYME CORP.

PA Borgford T;

XX WPI; 1999-009431/01.

PT New nucleic acid encoding ricin-like toxin with an interchain linker
PT cleaved by protease - is specific for diseased cells, useful for,
PT e.g. killing selectively cancer or infected cells
PS Claim 24; Fig 41D; 352pp; English.
XX

CC The present invention describes new purified and isolated nucleic acids
CC (I) encoding: (1) the A and B chains of a ricin-like toxin (II); and
CC (11) a heterologous linker, joining the two chains and including a
CC cleavage recognition site for a disease-specific protease (III). Also
CC described are: (1) plasmids or baculovirus transfer vectors that contain
CC (1); and (2) recombinant protein (IV) consisting of the A and B chains
CC of (II) joined by the specified linker. (IV), produced by expression of
CC (1) in host cells, are used to inhibit or kill diseased cells that
CC produce (III), particularly for treating cancers (e.g. leucocyte
CC proliferation; cancer of ovary, pancreas, breast or prostate; glioma) or
CC infections caused by fungi, parasites (e.g. malaria) or viruses (e.g.
CC cytomegalovirus (CMV), herpes, hepatitis, rhinovirus, laryngeotracheitis,
CC poliomyelitis or varicella zoster), also cystic fibrosis and multiple
CC sclerosis. Alternatively, (1) is used to express (IV) in vivo. (IV) is
CC toxic specifically for (III)-expressing cells and does not depend for
CC specificity on a cell-binding component. When used to treat virus-
CC infected cells, transcytosis and cytotoxicity of (IV) are increased by
CC retrograde translocation from endoplasmic reticulum to cytoplasm (which
CC some viruses exploit to avoid immune detection), so selectivity and
CC safety are further improved. (IV) are not toxic until chain A is
CC released and this occurs only in target cells. The present sequence
CC represents a specifically claimed mutant propeptidic linker from the
CC present invention.

CC Sequence 31 AA;

Query Match 100.0%; Score 21; DB 20; Length 31;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
1111
DB 20 HARL 23

RESULT 31
AAB34676

ID AAB34676 standard; Protein; 32 AA.

AC AAB34676;

DT 26-JAN-2001 (first entry)

DE Human secreted protein sequence encoded by gene 44 SEQ ID NO:160.

KW Human; secreted protein; diagnosis; immunosuppressive; antiarthritic;
KW antineumatic; antiproliferative; cytostatic; cardiac; vasotropic;
KW cerebroprotective; neurotropic; neuroprotective; antibacterial; virucide;
KW fungicide; ophthalmological; gene therapy; autoimmune disease; neoplasm;
KW hyperproliferative disorder; cancer; cardiovascular disorder; infection;
KW cerebrovascular disorder; angiogenesis; nervous system disorder;
KW ocular disorder; wound healing; skin aging; food additive; preservative.

OS Homo sapiens.

PN WO200056751-A1.

PD 28-SEP-2000.

PF 09-MAR-2000; 2000WO-US06013.

XX 19-MAR-1999; 99US-0125360.

PR 11-JUN-1999; 99US-0138626.

PR 03-DEC-1999; 99US-0168662.

XX (HUMA-) HUMAN GENOME SCI INC.

PA Rosen CA, Ruben SM, Komatsoulis G;

PT Isolated nucleic acid molecule encoding a human secreted protein is
PT used in preventing, treating or ameliorating a medical condition -

XX PS Disclosure; Page 413; 419pp; English.

CC The polynucleotide sequences given in AAC59738 to AAC59787 encode the
CC human secreted proteins given in AAB34577 to AAB34626. AAB34627 to
CC AAB34686 represent human secreted polypeptide sequences and proteins
CC homologous to them, which are given in the exemplification of the present
CC invention. Human secreted proteins have activities based on the tissues
CC and cells the genes are expressed in. Example of activities include:
CC antithalitic; immunosuppressive; antineumatic; antiproliferative;
CC cytosolic; cardiant; vasotropic; cerebroprotective; noctropic;
CC neuroprotective; antibacterial; virucide; fungicide; and
CC ophthalmological. The polynucleotides and proteins can be used to
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also
CC used in diagnosing a pathological condition or susceptibility to a
CC pathological condition. Disorders which are diagnosed or treated include
CC autoimmune diseases, hyperproliferative disorders e.g. neoplasms and
CC cancers of the breast or liver, cardiovascular disorders,
CC cerebrovascular disorders, angiogenesis, nervous system disorders,
CC infections caused by bacteria, viruses and fungi and ocular disorders.
CC The proteins can also be used to aid wound healing and epithelial cell
CC proliferation, to prevent skin aging due to sunburn, to maintain organs
CC before transplantation, for supporting cell culture of primary tissues,
CC to regenerate tissues and in chemotaxis. The proteins can also be used
CC as a food additive or preservative to increase or decrease storage
CC capabilities. AAC59729 to AAC59737 and AAB34576 represent sequences used
CC in the exemplification of the present invention.

XX SQ Sequence 32 AA;

Query Match 100.0%; Score 21; DB 21; Length 32;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
| | | |
DB 10 HARL 13

RESULT 32

ABBB1475
ID ABB31475 standard; Peptide; 32 AA.

XX AC ABB31475;

XX DT 01-FEB-2002 (first entry)

XX DE Peptide #4126 encoded by breast cell single exon nucleic acid probe.

XX KM Human; microarray; single exon probe; gene expression; breast;
XX disease; cancer.

XX OS Homo sapiens.

XX PN WO200157271-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US00662.

XX PR 04-FEB-2000; 2000US-0180312.

XX PR 26-MAY-2000; 2000US-0207456.

XX PR 30-JUN-2000; 2000US-0608408.

XX PR 03-AUG-2000; 2000US-0632366.

XX PR 21-SEP-2000; 2000US-0234687.

XX PR 27-SEP-2000; 2000US-0236359.

XX PR 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-496933/54.

XX PR New spatially-addressable set of single exon nucleic acid probes,
XX useful for measuring gene expression in sample derived from human
XX breast, comprises number of single exon nucleic acid probes
PS Claim 27; SEQ ID NO 14443; 327pp + sequence listing; English.

XX The invention relates to a spatially-addressable set of single exon
XX nucleic acid probes for measuring gene expression in a sample derived
XX from human breast and BT 474 cells. The method involves contacting
XX the probes with a collection of detectably labelled nucleic acids
XX derived from mRNA of human breast, and then measuring the label
XX bound to each probe of the microarray. The probes are useful for
XX verifying the expression of regions of genomic DNA predicted to
XX encode proteins. They are useful for gene discovery, and for
XX determining predisposition and/or prognosis breast disease. Gene
XX expression analysis is useful for assessing the toxicity of chemical
XX agents on cells. The microarray of this invention presents a far greater
XX diversity of probes for measuring gene expression, with far less bias
XX than expressed sequence tag microarrays. The method is suitable for
XX rapid production of functional information from genomic sequence. The
XX present sequence is a peptide encoded by a single exon nucleic acid
XX probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
XX from WPI at ftp.wpi.int/pub/published_pct_sequences.

XX SQ Sequence 32 AA;

Query Match 100.0%; Score 21; DB 22; Length 32;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
| | | |
DB 12 HARL 15

RESULT 33

ABBB3687
ID ABB36687 standard; Peptide; 32 AA.

XX AC ABB36687;

XX DT 04-FEB-2002 (first entry)

XX DE Peptide #4193 encoded by human foetal liver single exon probe.

XX KM Human; foetal liver; gene expression; single exon nucleic acid probe.

XX OS Homo sapiens.

XX PN WO200157277-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US00669.

XX PR 04-FEB-2000; 2000US-0180312.

XX PR 26-MAY-2000; 2000US-0207456.

XX PR 30-JUN-2000; 2000US-0608408.

XX PR 03-AUG-2000; 2000US-0632366.

XX PR 21-SEP-2000; 2000US-0234687.

XX PR 27-SEP-2000; 2000US-0236359.

XX PR 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-483447/52.

PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human fetal liver -
PS Claim 27; SEQ ID NO 29322; 639pp + sequence listing; English.
XX
CC The invention relates to a single exon nucleic acid probe for
CC measuring human gene expression in a sample derived from human foetal
CC liver. The single exon nucleic acid probes may be used for predicting,
CC measuring and displaying gene expression in samples derived from human
CC fetal liver. The present sequence is a peptide encoded by a single exon
CC nucleic acid probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
SQ
XX Sequence 32 AA;
Query Match 100.0%; Score 21; DB 22; Length 32;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 HARL 4
DB 12 HARL 15
RESULT 34
ABB22024
ID ABB22024 standard; Protein; 32 AA.
XX
AC ABB22024;
XX
DT 23-JAN-2002 (first entry)
XX
DE Protein #4023 encoded by probe for measuring heart cell gene expression.
XX
XX Human: gene expression; heart; microarray; vascular system;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
KW congenital heart disease.
XX
OS Homo sapiens.
XX
PN MO200157274-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00666.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-488899/53.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT hearts -
XX
PS Claim 15; SEQ ID NO 23794; 530pp; English.
XX
CC The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart (see
CC ABA21535-ABA41305). The present sequence is a protein encoded by one such
CC probe. The probes may be used for predicting, measuring and displaying
CC gene expression in samples derived from the human heart via microarrays.
CC By measuring gene expression, the probes are useful for predicting,

CC diagnosing, grading, staging, monitoring and prognosing diseases of the
CC human heart and vascular system e.g. cardiovascular disease.
CC hypertension, cardiac arrhythmias and congenital heart disease.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
SQ
XX Sequence 32 AA;
Query Match 100.0%; Score 21; DB 22; Length 32;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 HARL 4
DB 12 HARL 15
RESULT 35
AAM57451
ID AAM57451 standard; Protein; 32 AA.
XX
AC AAM57451;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe encoded protein SEQ ID NO: 29556.
XX
XX Human; brain expressed exon; gene expression analysis; probe;
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW epilepsy; cancer.
XX
OS Homo sapiens.
XX
PN WO200157275-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00667.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-483446/52.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT brains -
XX
XX Example 4; SEQ ID NO: 29556; 650pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is a protein encoded by one of
CC the probes of the invention.
XX
SQ Sequence 32 AA;
Query Match 100.0%; Score 21; DB 22; Length 32;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
|||
DB 12 HARL 15

RESULT 36
AAM69846
ID AAM69846 standard; Protein; 32 AA.

XX AAM69846;

XX AC AAM69846;

XX 06-NOV-2001 (first entry)

DE Human bone marrow expressed probe encoded protein SEQ ID NO: 30152.

KW Human; bone marrow expressed exon; gene expression analysis; probe;

KW microarray; cancer; leukemia; lymphoma; myeloma.

OS Homo sapiens.

XX WO200157276-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00668.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207456.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-0234687.

XX 27-SEP-2000; 2000US-0236359.

XX 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488900/53.

XX Human genome-derived single exon nucleic acid probes useful for

XX analyzing gene expression in human bone marrow -

XX Example 4; SEQ ID NO: 30152; 658bp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid

XX probes which are derived from genomic sequences expressed in the human

XX bone marrow. They can be used to measure gene expression in bone marrow

XX samples, which may enable the improved diagnosis and treatment of cancers

XX such as lymphoma, leukemia and myeloma. The present sequence is a

XX protein encoded by one of the probes of the invention.

XX Sequence 32 AA;

XX Query Match 100.0%; Score 21; DB 22; Length 32;

XX Best Local Similarity 100.0%; Pred. No. 2.4e+02;

XX Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

KW Probe; human; microarray; gene expression; cervical epithelial cell;

KW cervical cancer.

XX Homo sapiens.

XX WO200157278-A2.

XX 30-JAN-2001; 2001WO-US00670.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207456.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-0234687.

XX 27-SEP-2000; 2000US-0236359.

XX 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488901/53.

XX Human genome-derived single exon nucleic acid probes useful for

XX analyzing gene expression in human cervical epithelial cells -

XX Claim 27; SEQ ID NO 22494; 487bp; English.

XX The present invention relates to human single exon nucleic acid probes

XX (SENP; see AAI10068-AA128459). The present sequence is a peptide encoded

XX by one such probe. The SENPs are derived from human HeLa cells. The SENPs

XX can be used to produce a single exon microarray, which can be used for

XX measuring human gene expression in a sample derived from human cervical

XX epithelial cells. By measuring gene expression, the probes are therefore

XX useful in grading and/or staging of diseases of the cervix, notably

XX cervical cancer.

XX Note: The sequence data for this patent did not form part of the printed

XX specification, but was obtained in electronic format directly from WIPO

XX at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 32 AA;

XX Query Match 100.0%; Score 21; DB 22; Length 32;

XX Best Local Similarity 100.0%; Pred. No. 2.4e+02;

XX Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX 1 HARL 4

XX 12 HARL 15

XX 17-OCT-2001 (first entry)

XX Peptide #4222 encoded by probe for measuring placental gene expression.

XX Probe; microarray; human; placenta; antenatal diagnosis;

XX genetic disorder.

XX Homo sapiens.

XX WO200157272-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00663.

```
xx 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0633366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
xx
PA (MOLE-) MOLECULAR DYNAMICS INC.
xx
PI Penn SG, Hanzel DK, Chen W, Rank DR;
xx
DR WPI; 2001-488897/53.
xx
PT Human genome-derived single exon nucleic acid probes useful for
xx analyzing gene expression in human placenta -
xx
PS Claim 27; SEQ ID No 30454; 654bp; English.
xx
CC The present invention relates to single exon nucleic acid probes (SENP:
CC see AI1315-AI157546). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for producing a microarray for
CC predicting, measuring and displaying gene expression in samples derived
CC from human placenta. The probes are useful for antenatal diagnosis of
CC human genetic disorders.
xx
SQ Sequence 32 AA;
xx
Query Match 100.0%; Score 21; DB 22; Length 32;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 HARL 4
DB 12 HARL 15
xx
RESULT 39
AAM05331
ID AAM05331 standard; Protein; 32 AA.
xx
AC AAM05331;
xx
DT 09-OCT-2001 (first entry)
xx
DE Peptide #4013 encoded by probe for measuring breast gene expression.
xx
KW Probe; human; breast disease; breast cancer; development disorder;
KW inflammatory disease; proliferative breast disease; non-carcinoma tumour.
xx
OS Homo sapiens.
xx
PN WO200157270-A2.
xx
PD 09-AUG-2001.
xx
PF 29-JAN-2001; 2001WO-US00661.
xx
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0633366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
xx
PA (MOLE-) MOLECULAR DYNAMICS INC.
xx
PI Penn SG, Hanzel DK, Chen W, Rank DR;
xx
DR WPI; 2001-476286/51.
xx
```

```
PT Novel single exon nucleic acid probe used to measuring gene expression
PT in a human breast -
xx
PS Claim 27; SEQ ID No 14071; 322pp; English.
xx
CC The present invention relates to novel single exon nucleic acid probes
CC (see AI100010-AI110067). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for measuring human gene expression in
CC a human breast sample, where the probe hybridises at high stringency to a
CC nucleic acid expressed in the human breast. The probes are useful for
CC predicting, diagnosing, grading, staging, monitoring and prognosing
CC diseases of the human breast, particularly those diseases with polygenic
CC aetiology. The diseases include: breast cancer, disorders of development,
CC inflammatory diseases of the breast, fibrocystic changes, proliferative
CC breast disease and non-carcinoma tumours.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
xx
SQ Sequence 32 AA;
xx
Query Match 100.0%; Score 21; DB 22; Length 32;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 HARL 4
DB 12 HARL 15
xx
RESULT 40
ABG39474
ID ABG39474 standard; Peptide; 32 AA.
xx
AC ABG39474;
xx
DT 19-AUG-2002 (first entry)
xx
DE Human peptide encoded by genome-derived single exon probe SEQ ID 29139.
xx
KW Human; single exon probe; asthma; lung cancer; COPD; IHD;
KW chronic obstructive pulmonary disease; interstitial lung disease;
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW pulmonary histiocytosis; lymphangioleiomyomatosis; Kargener syndrome;
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW primary ciliary dyskinesia; pulmonary hypertension;
KW hyaline membrane disease.
xx
OS Homo sapiens.
xx
PN WO200186603-A2.
xx
PD 15-NOV-2001.
xx
PF 30-JAN-2001; 2001WO-US00665.
xx
PR 04-FEB-2000; 2000US-180312P.
PR 26-MAY-2000; 2000US-207456P.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0633366.
PR 21-SEP-2000; 2000US-234687P.
PR 27-SEP-2000; 2000US-236359P.
PR 04-OCT-2000; 2000GB-0024263.
xx
PA (MOLE-) MOLECULAR DYNAMICS INC.
xx
PI Penn SG, Hanzel DK, Chen W, Rank DR;
xx
DR WPI; 2002-114183/15.
xx
PT Spatially-addressable set of single exon nucleic acid probes, used to
```

PT measure gene expression in human lung samples -
 XX Claim 27; SEQ ID NO 29139; 634pp; English.
 XX
 CC The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human lung comprising single exon nucleic acid probes having one of
 CC 12614 nucleic acid sequences mentioned in the specification, or their
 CC complements or the 12387 open reading frames derived from the 12614
 CC probes. Also included are a microarray comprising the novel set of
 CC probes; the novel set of probes which hybridise at high stringency to a
 CC nucleic acid expressed in the human lung; measuring gene expression in a
 CC sample derived from human lung, comprising (a) contacting the array with
 CC a collection of detectably labeled nucleic acids derived from human lung
 CC mRNA, and (b) measuring the label detectably bound to each probe of
 CC the array; identifying exons in a eukaryotic genome, comprising
 CC (a) algorithmically predicting at least one exon from genomic sequences
 CC of the eukaryote; and (b) detecting specific hybridisation of detectably
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridisation to a single exon
 CC microarrays having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned to a single gene; a peptide comprising one
 CC of 12011 sequences, mentioned in the specification, or encoded by the
 CC probes/open reading frames (ORF). The probes are used for gene
 CC expression analysis, and for identifying exons in a gene, particularly
 CC using human lung derived mRNA and for the study of lung diseases
 CC such as asthma, lung cancer, chronic obstructive pulmonary disease
 CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
 CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
 CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
 CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
 CC pulmonary alveolar proteinosis, Karagenen syndrome, fibrocystic
 CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
 CC and hyaline membrane disease. The present sequence is a peptide/protein
 CC encoded by a single exon probe of the invention.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX
 SQ Sequence 32 AA;
 Query Match 100.0%; Score 21; DB 23; Length 32;
 Best Local Similarity 100.0%; Pred. No. 2.4e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 HARL 4
 DB 12 HARL 15
 ||||
 RESULT 41
 AAO09129
 ID AAO09129 standard; Protein; 36 AA.
 XX
 AC AAO09129;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human polypeptide SEQ ID NO 23021.
 XX
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorders; arthritis; inflammation.
 XX
 OS Homo sapiens.
 XX

PN W0200164835-A2.
 XX
 PD 07-SEP-2001.
 XX
 XX 26-FEB-2001; 2001WO-US04927.
 PF
 XX 26-FEB-2000; 2000US-0515126.
 PR 18-MAY-2000; 2000US-0577409.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT;
 XX
 DR WPI; 2001-514838/56.
 DR N-PSDB; AAI89060.
 XX
 PT Isolated nucleic acids and polypeptides, useful for preventing
 PT diagnosing and treating e.g. leukaemia, inflammation and immune
 PT disorders -
 PS Claim 20; SEQ ID NO 23021; 1399pp + Sequence Listing; English.
 XX
 CC The invention relates to human polynucleotides (AAI79941-AAI93841) and
 CC the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX
 SQ Sequence 36 AA;
 Query Match 100.0%; Score 21; DB 22; Length 36;
 Best Local Similarity 100.0%; Pred. No. 2.7e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 HARL 4
 DB 13 HARL 16
 ||||
 RESULT 42
 AAO08247
 ID AAO08247 standard; Protein; 37 AA.
 XX
 AC AAO08247;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human polypeptide SEQ ID NO 22139.
 XX
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorders; arthritis; inflammation.
 XX
 OS Homo sapiens.
 XX
 PN W0200164835-A2.
 PN
 XX 07-SEP-2001.
 PD
 XX 26-FEB-2001; 2001WO-US04927.
 PF
 XX 26-FEB-2000; 2000US-0515126.
 PR 18-MAY-2000; 2000US-0577409.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT;
 XX
 DR WPI; 2001-514838/56.
 DR N-PSDB; AAI89060.
 XX
 PT Isolated nucleic acids and polypeptides, useful for preventing
 PT diagnosing and treating e.g. leukaemia, inflammation and immune
 PT disorders -
 PS Claim 20; SEQ ID NO 23021; 1399pp + Sequence Listing; English.
 XX
 CC The invention relates to human polynucleotides (AAI79941-AAI93841) and
 CC the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX
 SQ Sequence 36 AA;
 Query Match 100.0%; Score 21; DB 22; Length 36;
 Best Local Similarity 100.0%; Pred. No. 2.7e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 HARL 4
 DB 13 HARL 16
 ||||
 RESULT 42
 AAO08247
 ID AAO08247 standard; Protein; 37 AA.
 XX
 AC AAO08247;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human polypeptide SEQ ID NO 22139.
 XX
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorders; arthritis; inflammation.
 XX
 OS Homo sapiens.
 XX
 PN W0200164835-A2.
 PN
 XX 07-SEP-2001.
 PD
 XX 26-FEB-2001; 2001WO-US04927.
 PF
 XX 26-FEB-2000; 2000US-0515126.
 PR 18-MAY-2000; 2000US-0577409.
 XX

XX (HYSE-) HYSEQ INC.
PA Tang YT, Liu C, Drmanac RF;
PI WPI; 2001-514838/56.
DR N-PSDB; AA188178.
XX
PT Isolated nucleic acids and polypeptides, useful for preventing
PT diagnosing and treating e.g. leukaemia, inflammation and immune
PT disorders -
PS
XX Claim 20; SEQ ID NO 22139; 1399pp + Sequence listing; English.
XX
CC The invention relates to human polynucleotides (AA179941-AA193841) and
CC the encoded proteins (AA000010-AA013910) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, hematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SO Sequence 37 AA;

Query Match 100.0%; Score 21; DB 22; Length 37;
Best Local Similarity 100.0%; Pred. No. 2,7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
DB 34 HARL 37

RESULT 43
AAM94835
ID AAM94835 standard; Protein; 40 AA.
XX
AC AAM94835;
XX
DT 21-NOV-2001 (first entry)
XX
DE Human reproductive system related antigen SEQ ID NO: 3493.
XX
KW Human; reproductive system related antigen; reproductive system disorder;
KW cancer; gene therapy.
XX
OS Homo sapiens.
XX
PN WO200155320-A2.
PD
XX
PD 02-AUG-2001.
PF
XX 17-JAN-2001; 2001WO-US01339.
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0188874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR

PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229503.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239335.
PR 13-OCT-2000; 2000US-0239337.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.

PR	20-OCT-2000;	2000US-0241809.
PR	20-OCT-2000;	2000US-0241826.
PR	01-NOV-2000;	2000US-0244617.
PR	08-NOV-2000;	2000US-0246474.
PR	08-NOV-2000;	2000US-0246475.
PR	08-NOV-2000;	2000US-0246476.
PR	08-NOV-2000;	2000US-0246477.
PR	08-NOV-2000;	2000US-0246478.
PR	08-NOV-2000;	2000US-0246523.
PR	08-NOV-2000;	2000US-0246524.
PR	08-NOV-2000;	2000US-0246525.
PR	08-NOV-2000;	2000US-0246526.
PR	08-NOV-2000;	2000US-0246527.
PR	08-NOV-2000;	2000US-0246528.
PR	08-NOV-2000;	2000US-0246532.
PR	08-NOV-2000;	2000US-0246609.
PR	08-NOV-2000;	2000US-0246610.
PR	08-NOV-2000;	2000US-0246613.
PR	17-NOV-2000;	2000US-0249207.
PR	17-NOV-2000;	2000US-0249208.
PR	17-NOV-2000;	2000US-0249209.
PR	17-NOV-2000;	2000US-0249210.
PR	17-NOV-2000;	2000US-0249211.
PR	17-NOV-2000;	2000US-0249212.
PR	17-NOV-2000;	2000US-0249213.
PR	17-NOV-2000;	2000US-0249214.
PR	17-NOV-2000;	2000US-0249215.
PR	17-NOV-2000;	2000US-0249216.
PR	17-NOV-2000;	2000US-0249217.
PR	17-NOV-2000;	2000US-0249218.
PR	17-NOV-2000;	2000US-0249244.
PR	17-NOV-2000;	2000US-0249245.
PR	17-NOV-2000;	2000US-0249246.
PR	17-NOV-2000;	2000US-0249265.
PR	17-NOV-2000;	2000US-0249297.
PR	17-NOV-2000;	2000US-0249299.
PR	17-NOV-2000;	2000US-0249300.
PR	01-DEC-2000;	2000US-0250160.
PR	01-DEC-2000;	2000US-0250391.
PR	05-DEC-2000;	2000US-0251030.
PR	05-DEC-2000;	2000US-0251988.
PR	06-DEC-2000;	2000US-0256719.
PR	08-DEC-2000;	2000US-0251479.
PR	08-DEC-2000;	2000US-0251856.
PR	08-DEC-2000;	2000US-0251868.
PR	08-DEC-2000;	2000US-0251869.
PR	08-DEC-2000;	2000US-0251989.
PR	08-DEC-2000;	2000US-0251990.
PR	11-DEC-2000;	2000US-0254097.
PR	05-JAN-2001;	2001US-0259678.
PA	(HOMA-) HUMAN GENOME SCI INC.	
XX		
XI	Rosen CA, Barash SC, Ruben SM;	
XX	WPI; 2001-465570/50.	
DR	N-PSDB; AAL00805.	
DR		
PT	Isolated nucleic acid molecule encoding a reproductive system antigen	-
PT	is used in preventing, treating or ameliorating a medical condition	
XX		
XX	Claim 11; SEQ ID NO 3493; 1297Pp + Sequence Listing; English.	
CC	The present invention provides the protein and coding sequences of a	
CC	number of human reproductive system related antigens. These can be used	
CC	in the prevention and treatment of reproductive system disorders,	
CC	including cancer. The present sequence is a protein of the invention.	
XX		
SQ	Sequence 40 AA:	
Query Match	100.0%; Score 21; DB 22; Length 40;	
Best Local Similarity	100.0%; Pred. No. 3e+02;	

[illegible]

AAU20929
ID AAU20929 standard; Protein; 43 AA.
XX
AC AAU20929;
XX
DT 18-DEC-2001 (first entry)
XX
DE Human novel foetal antigen, SEQ ID NO 1173.
XX
KW Human; foetal tissue antigen; antiinflammatory; neuroprotective;
immunomodulator; cardiovascular; cytosolic; nephrothropic;
KW cardiovascular; autoimmune disease; rheumatoid arthritis;
KW hyperproliferative disorder; breast neoplasia; cancer;
KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
KW cerebral ischemia; angiogenesis; nervous system disorder;
KW Alzheimer's disease; infection; ocular disorder; corneal infection;
wound healing; epithelial cell proliferation; food additive.
XX
OS Homo sapiens.
XX
PN WO20015312-A2.
XX
PD 02-AUG-2001.
XX
PE 17-JAN-2001; 2001WO-US01321.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226686.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.

PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234423.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239937.
PR 13-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0244674.
PR 08-NOV-2000; 2000US-0244675.
PR 08-NOV-2000; 2000US-0244676.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.

PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 01-DEC-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Ruden SM;
 XX
 DR WPI; 2001-488782/53.
 DR N-PSDB; AAS33749.
 XX
 PT New polynucleotides and polypeptides for diagnosing, treating,
 PT preventing or prognosing e.g. diseases or disorders of the nervous,
 PT musculoskeletal, excretory, gastrointestinal, reproductive, and
 PT respiratory systems
 XX
 PS Claim 11; SEQ ID NO 1173; 642bp; English.
 XX
 CC The invention relates to novel nucleic acids encoding novel human foetal
 CC antigens. The nucleic acids and proteins are used to prevent, treat (e.g.
 CC by gene therapy) or ameliorate a medical condition in e.g. humans, mice,
 CC rabbits, goats, horses, cats, dogs, chickens or sheep. They
 CC are also used in diagnosing a pathological condition or susceptibility
 CC to a pathological condition. The antibodies to the antigens can also
 CC be used in alleviating symptoms associated with the disorders and in
 CC diagnostic immunoassays e.g. radioimmunoassays or enzyme linked
 CC immunosorbent assays (ELISA). Disorders which are diagnosed or treated
 CC include autoimmune diseases e.g. rheumatoid arthritis,
 CC hyperproliferative disorders e.g. neoplasms of the breast or liver,
 CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders
 CC e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g.
 CC Alzheimer's disease, infections caused by bacteria, viruses and fungi
 CC and ocular disorders e.g. corneal infection. The polypeptides can also
 CC be used to aid wound healing and epithelial cell proliferation, to
 CC prevent skin aging due to sunburn, to maintain organs before
 CC transplantation, for supporting cell culture of primary tissues, to
 CC regenerate tissues and in chemotaxis. The polypeptides can also be used
 CC as a food additive or preservative to increase or decrease storage
 CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,
 CC minerals, cofactors and other nutritional components. Numerous
 CC examples of diseases and disorders treated by the nucleic acids and
 CC proteins are given in the specification. The present sequence

Query Match 100.0%; Score 21; DB 22; Length 43;
 Best Local Similarity 100.0%; Pred. No. 3.2e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 HARL 4
 Db 11 HARL 14

RESULT 46
 AAR85922
 ID AAR85922 standard; Peptide; 47 AA.
 XX
 AC AAR85922;
 XX
 DT 14-FEB-1996 (first entry)
 XX

DE Protein tyrosine-kinase SAL-S1 fragment.
 XX
 KW Protein tyrosine-kinase; SAL-S1; agonist; cell growth;
 KW differentiation.
 XX
 OS Homo sapiens.
 XX
 PN W09527061-A1.
 XX
 PD 12-OCT-1995.
 XX
 PF 04-APR-1995; 95WO-US04228.
 XX
 PR 04-APR-1994; 94US-0222616.
 XX
 PA (GETH) GENENTECH INC.
 XX
 PI Bennett BD, Goeddel D, Lee JM, Matthews W, Tsai SP,
 PI Wood WI;
 XX
 DR WPI; 1995-366160/47.
 DR N-PSDB; AAT03089.
 XX
 PT Agonist antibodies which activate specific protein tyrosine
 PT kinase(s) - also activate chimeric proteins of kinase extracellular
 PT domain and Ig constant domain, useful for studying, and therapeutic
 PT modulation of, cell growth and differentiation
 XX
 PS Disclosure; Page 35; 125pp; English.
 XX
 CC DNA probes based on protein tyrosine-kinase (PTK) sequences were used
 CC to screen cDNA libraries to identify novel PTK genes. A SAL-S1 gene
 CC fragment (see AAT03089) isolated from several megakaryocytic cell lines
 CC encoded a peptide having the sequence given in AAR85922 (see AAR85937
 CC for the full-length sequence). The SAL-S1 protein shows sequence
 CC homology with Fln/Frk family PTKs.
 XX
 SQ Sequence 47 AA;
 OY 1 HARL 4
 Db 35 HARL 38

Query Match 100.0%; Score 21; DB 16; Length 47;
 Best Local Similarity 100.0%; Pred. No. 3.5e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 HARL 4
 Db 35 HARL 38

RESULT 47
 AAB51456
 ID AAB51456 standard; Protein; 47 AA.
 XX
 AC AAB51456;
 XX
 DT 16-FEB-2001 (first entry)
 XX
 DE Human secreted protein BIAST search protein SEQ ID NO: 133.
 XX
 KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
 KW vulnery; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein.
 XX
 OS Homo sapiens.
 XX
 PN W0200058495-A1.
 XX
 PD 05-OCT-2000.
 XX
 PR 23-MAR-2000; 2000WO-US07661.
 XX
 DT 26-MAR-1999; 99US-0126504.
 XX

XX 07-JAN-2000; 2000US-0174847.
XX
XA (HDM-A) HDMAN GENOME SCI INC.
XX
XZ Rosen CA, Ruben SM, Komatsoulis G;
XX
DR WPI: 2000-611720/58.
XX
PT New nucleic acid molecules encoding 45 human secreted proteins for
PT diagnosing, preventing, treating or ameliorating medical conditions and
PT used as food additives or preservatives -
XX
XX
PS Disclosure; Page 385; 410pp; English.
XX
CC The invention relates to the isolation of genes AC93310-C93354 encoding
CC the human secreted proteins AAB51380-B1413. The genes and proteins are
CC useful for preventing, ameliorating or treating medical conditions, e.g.
CC by protein or gene therapy. The genes are isolated from a range of human
CC tissues disclosed in the specification. The nucleic acids, proteins,
CC antibodies and (ant)agonists are useful in the diagnosis, treatment and
CC prevention of: (a) cancer, e.g. breast and ovarian cancer, and other
CC cancers of the adrenal gland, bone, bone marrow, breast,
CC gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders
CC e.g. Addison's disease, allergies, autoimmune haemolytic anaemia,
CC autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple
CC sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC cardiovascular disorders such as myocardial ischaemias; (d) wound
CC healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy;
CC and (f) infectious diseases such as viral, bacterial, fungal and
CC parasitic infections. The present sequence is a protein isolated in the
CC present invention.
XX
XX Sequence 47 AA;
XQ

Query Match	100.0 %	Score 21	DB 21	Length 47
Best Local Similarity	100.0 %	Pred. No. 3	Se+02	
Matches 4, Conservative	0	Mismatches 0	Indels 0	Gaps 0

QY	1	HARL	4
Db	38	HARL	41

RESULT 48
AAB51460
ID AAB51460 standard; Protein; 47 AA

AC	AAB51460;
XX	
DT	16-FEB-2001 (first entry)

Human secreted protein BLAST search SEQ ID NO: 137

KM Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KM antiallergic; hepatotropic; antidiabetic; antinflammatory; antitumor;
 KM vulnery; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KM cardiac; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KM neurological disease; infection; human; secreted protein.

OS Homo sapiens.

PN W0200058495-A1.

PD 05-OCT-2000.

PF 23-MAR-2000; 2000WO-US07661.

PR 26-MAR-1999; 99US-0126504.

XX
PA (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Ruben SM, Komatsoulls G;

XX
DR WPI; 2000-611720/58.
XX
XX New nucleic acid molecules encoding 45 human secreted proteins for
PT diagnosing, preventing, treating or ameliorating medical conditions and
PT used as food additives or preservatives -
XX
XX Disclosure; Page 386; 410pp; English.

CC The invention relates to the isolation of genes AC933310-093354 encoding
CC the human secreted proteins AAB1380-B51423. The genes and proteins are
CC useful for preventing, ameliorating or treating medical conditions, e.g.
CC by protein or gene therapy. The genes are isolated from a range of human
CC tissues disclosed in the specification. The nucleic acids, proteins,
CC antibodies and (ant)agonists are useful in the diagnosis, treatment and
CC prevention of: (a) cancer, e.g. breast and ovarian cancer, and other
CC cancers of the adrenal gland, bone, bone marrow, breast,
CC gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders
CC e.g. Addison's disease, allergies, autoimmune hemolytic anaemia,
CC autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple
CC sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC cardiovascular disorders such as myocardial ischaemias; (d) wound
CC healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy;
CC and (f) infectious diseases such as viral, bacterial, fungal and
CC parasitic infections. The present sequence is a protein isolated in the
CC present invention.

.....	SQ	Sequence	47 AA;
-------	----	----------	--------

Query Match	100.0%;	Score 21;	DB 21;	Length 47;
Best Local Similarity	100.0%;	Pred. No. 3.5e+02;		
Matches 4; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0

QY	1 HARL 4
Db	38 HARL 41

RESULT 49
AAB51688
ID AAB51688 standard; Protein; 47 AA.

AC AAB51688;
XX
DT 26-FEB-2001 (first entry)

DE Human secreted protein sequence encoded by gene 23 SEQ ID NO:128.

KM Human; secreted protein; diagnosis; immunosuppressive; antitubercular;
 KM antirheumatic; antiproliferative; cytostatic; candida; vasotropic;
 KM cardioprotective; nontropic; neuroprotective; antibacterial; vinicid;
 KM fungicide; ophthalmological; gene therapy; autoimmune disease; infection
 KM hyperproliferative disorder; cardiovascular disorder; angiogenesis;
 KM cerebrovascular disorder; nervous system disorder; ocular disorder;
 KM wound healing; skin aging; food additive; preservative.

OS Homo sapiens.

PN W0200061620-A1.

PD 19-OCT-2000.

PF 06-APR-2000; 2000WO-US09069.

PR 09-APR-1999; 99US-0128702.

[illegible]

PA (ROSE/) ROSEN C A.

PI Rosen CA, Ruben SM, Komatsoulis G,

DR WPI; 2000-619225/59.

XX Isolated nucleic acid molecule encoding a human secreted protein is
PT used in preventing, treating or ameliorating a medical condition
XX
XX
PS Disclosure; Page 504; 540pp; English.

CC The polynucleotide sequences given in AAC93364 to AAC93412 encode the
CC human secreted proteins given in AAB51620 to AAB51668. AAB51669 to
CC AAB51722 represent human secreted polypeptide sequences and proteins
CC homologous to them, which are given in the exemplification of the present
CC invention. Human secreted proteins have activities based on the tissues
CC and cells the genes are expressed in. Examples of activities include:
CC antitumor; immunosuppressive; antirheumatic; antiproliferative;
CC cytostatic; cardiant; vasotropic; cerebroprotective; neurotropic;
CC neuroprotective; antibacterial; virucide; fungicide; and
CC ophthalmological. The polynucleotides and polypeptides can be used to
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
CC in diagnosing a pathological condition or susceptibility to a
CC pathological condition. Disorders which are diagnosed or treated include
CC autoimmune diseases, hyperproliferative disorders, cardiovascular
CC disorders, cerebrovascular disorders, angiogenesis, nervous system
CC disorders, infections caused by bacteria, viruses and fungi and ocular
CC disorders. The polypeptides can also be used to aid wound healing and
CC epithelial cell proliferation, to prevent skin aging due to sunburn, to
CC maintain organs before transplantation, for supporting cell culture of
CC primary tissues, to regenerate tissues and in chemotaxis. The
CC polypeptides can also be used as a food additive or preservative to
CC increase or decrease storage capabilities. AAC93355 to AAC93363 and
CC AAB51619 represent sequences which are used in the exemplification of the
CC present invention.

SQ Sequence 47 AA;

Query Match 100.0%; Score 21; DB 21; Length 47;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
|||
Db 38 HARL 41

RESULT 50

AA011479

ID AA011479 standard; Protein; 47 AA.

XX AA011479;

AC AA011479;

XX 06-NOV-2001 (first entry)

DE Human polypeptide SEQ ID NO 25371.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;

KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;

KW tissue growth factor; immunomodulatory; cancer; leukaemia;

KW nervous system disorders; arthritis; inflammation.

XX Homo sapiens.

OS WO200164835-A2.

PN 07-SEP-2001.

PD 26-FEB-2001; 2001WO-0504927.

XX 28-FEB-2000; 2000US-0515126.

PR 18-MAY-2000; 2000US-0577409.

XX (HYSE-) HYSEQ INC.

PA Tang YT, Liu C, Drmanac RT;

XX

DR WPI: 2001-514838/56.

DR N-PSDB; AAI91410.

XX Isolated nucleic acids and polypeptides, useful for preventing
PT diagnosing and treating e.g. leukaemia, inflammation and immune
PT disorders -

PS Claim 20; SEQ ID NO 25371; 1399pp + Sequence Listing; English.

XX The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.

CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 47 AA;

Query Match 100.0%; Score 21; DB 22; Length 47;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
|||
Db 37 HARL 40

Search completed: November 13, 2002, 13:41:00
Job time : 37 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2002, 13:39:28 ; Search time 16 Seconds
(without alignments)
24.034 Million cell updates/sec

Title: US-09-697-590-2_COPY_292_295

Perfect score: 21

Sequence: 1 HARL 4

Scoring table: BLOSUM62
Gap 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

PIR_73: *
1: p1r1: *
2: p1r2: *
3: p1r3: *
4: p1r4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	21	100.0	44	2	B97644
2	21	100.0	59	2	B69101
3	21	100.0	65	2	S28481
4	21	100.0	67	2	AF2590
5	21	100.0	71	2	B69337
6	21	100.0	74	2	C70501
7	21	100.0	75	2	T06193
8	21	100.0	80	2	F82684
9	21	100.0	83	2	G84144
10	21	100.0	85	2	H97404
11	21	100.0	85	2	AA2622
12	21	100.0	93	2	E83358
13	21	100.0	93	2	T20227
14	21	100.0	96	2	S38916
15	21	100.0	100	2	A97833
16	21	100.0	101	2	C81012
17	21	100.0	103	2	A97484
18	21	100.0	103	2	A12701
19	21	100.0	103	2	D07412
20	21	100.0	106	2	S04786
21	21	100.0	107	2	U00389
22	21	100.0	107	2	T07540
23	21	100.0	108	2	S43154
24	21	100.0	116	1	OOCV81
25	21	100.0	116	2	D49339
26	21	100.0	118	2	I57506
27	21	100.0	118	2	I77386
28	21	100.0	119	2	E70718
29	21	100.0	120	2	C96008

30	21	100.0	121	2	E87139	50S ribosomal prot
31	21	100.0	121	2	T37737	hypothetical prote
32	21	100.0	122	2	T45382	ribosomal protein
33	21	100.0	122	2	G81198	hypothetical prote
34	21	100.0	122	2	C81775	hypothetical prote
35	21	100.0	124	2	D85843	unknown protein en
36	21	100.0	124	2	D85630	probable tail asse
37	21	100.0	124	2	F90822	probable tail asse
38	21	100.0	124	2	A90854	probable tail asse
39	21	100.0	124	2	B90909	probable tail asse
40	21	100.0	124	2	E90876	probable tail asse
41	21	100.0	125	2	T49512	hypothetical prote
42	21	100.0	127	2	PC2265	cytochrome P450 pr
43	21	100.0	127	2	B91218	thioredoxin 1 (imp
44	21	100.0	127	2	C86064	thioredoxin 1 (imp
45	21	100.0	129	2	G72460	hypothetical prote

ALIGNMENTS

RESULT 1
B97644
hypothetical protein AGR_C_4298 [imported] - Agrobacterium tumefaciens (strain C58,
C:Species: Agrobacterium tumefaciens
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 11-Jan-2002
C:Accession: B97644
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Ouncillo, B.; Gold
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughy, D.; Scott, C.; Lappas, C.; Markelz,
Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium
A:Reference number: A97359; PMID:11743194
A:Accession: B97644
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-44 <KUR>
A:Cross-references: GB:AE007869; PIDD:AAK8107.1; PID:g15157539; GSPDB:GN00169
C:Genetics:
A:Gene: AGR_C_4298
A:Map position: circular chromosome

Query Match 100.0%; Score 21; DB 2; Length 44;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HARL 4
DB 12 HARL 15

RESULT 2
B69101
hypothetical protein MTH1750 - Methanobacterium thermoautotrophicum (strain Delta H)
C:Species: Methanobacterium thermoautotrophicum
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
C:Accession: B69101
R:Smith, D.R.; Doncette-Stamm, L.A.; Delonghery, C.; Lee, R.; Dubois, J.; Aldredge,
; Liu, D.; Spadafora, R.; Vitale, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani,
K.; Church, G.M.; Daniels, C.J.; Mo, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: f
A:Reference number: A69000; MUID:98037514; PMID:9371463
A:Accession: B69101
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-59 <MTH>
A:Cross-references: GB:AE000930; GB:AE000666; NID:g2622872; PIDD:AAB86220.1; PID:g26
A:Experimental source: strain Delta H
C:Genetics:
A:Gene: MTH1750

Query Match 100.0%; Score 21; DB 2; Length 59;
Best Local Similarity 100.0%; Pred. No. 93;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HARL 4
Db 15 HARL 18

RESULT 3

rfbQ protein - *Vibrio cholerae* (strain O17)
C:Species: *Vibrio cholerae*
A:Variety: strain O17
C:Date: 12-Mar-1993 #sequence_revision 12-Mar-1993 #text_change 08-Oct-1999
C:Accession: S28481
R:Manning, P. A.
Submitted to the EMBL Data Library, May 1991
A:Reference number: S28467
A:Accession: S28481
A:Molecule type: DNA
A:Residues: 1-65 <MAN>
A:Cross-references: EMBL:X59554; NID:g48381; PIDN:CAA42147.1; PID:g48396
A:Experimental source: strain O17
C:Genetics:
A:Gene: rfbQ

Query Match 100.0%; Score 21; DB 2; Length 65;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HARL 4
Db 53 HARL 56

RESULT 4

hypothetical protein Atu0115 [imported] - *Agrobacterium tumefaciens* (strain C58, Dupont)
C:Species: *Agrobacterium tumefaciens*
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002
C:Accession: AF2590
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; WOO, I.
erage, G.; Giliat, W.; Grant, C.; Gunthner, D.; Kutyavln, T.; Levy, R.; Li, M.; Mclell
Science 294, 2317-2323, 2001
A:Authors: Ioo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.
A:Reference number: AB2577; PMID:11743193
A:Accession: AF2590
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-67 <KOR>
A:Cross-references: GB:AE008688; PIDN:AAL41140.1; PID:g17738435; GSPDB:GN00186
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: Atu0115
A:Map position: circular chromosome

Query Match 100.0%; Score 21; DB 2; Length 67;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HARL 4
Db 53 HARL 56

RESULT 5

B69337
hypothetical protein AF0698 - *Archaeoglobus fulgidus*
C:Species: *Archaeoglobus fulgidus*
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
C:Accession: B69337

R:Kienk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Doc
.: Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirschner, J
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Wetmann, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A:Authors: Uitterlinden, T.; Cotton, M.D.; Spriggs, T.; Artlich, P.; Kaine, B.P.; Sykes,
Smith, H.O.; Woese, C.R.; Venter, J.C.
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing arc
A:Reference number: A69250; MUID:98049343; PMID:9389475
A:Accession: B69337
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-71 <KLE>
A:Cross-references: GB:AE001056; GB:AE000782; NID:g2689379; PIDN:AAB90546.1; PID:g264

Query Match 100.0%; Score 21; DB 2; Length 71;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HARL 4
Db 4 HARL 7

RESULT 6

hypothetical protein RV1684 - *Mycobacterium tuberculosis* (strain H37RV)
C:Species: *Mycobacterium tuberculosis*
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C:Accession: C70501
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon
; Connor, R.; Davies, R.; Devlin, K.; Feldwell, T.; Gentles, S.; Hamlin, N.; Holroyd,
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete gen
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: C70501
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-74 <COL>
A:Cross-references: GB:298268; GB:AL123456; NID:g3261839; PIDN:CAB10941.1; PID:e33277
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: RV1684

Query Match 100.0%; Score 21; DB 2; Length 74;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HARL 4
Db 61 HARL 64

RESULT 7

T06193
chlorophyll a/b-binding protein - barley (fragment)
C:Species: *Hordeum vulgare* (barley)
C:Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 20-Jun-2000
C:Accession: T06193
R:Humbek, K.; Krupinska, K.
Submitted to the EMBL Data Library, May 1998
A:Description: Differential regulation of chlorophyll a/b-binding proteins and ELIP
A:Reference number: Z15522
A:Accession: T06193
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-75 <HUN>
A:Cross-references: EMBL:AU006296; PIDN:CAA06961.1
A:Experimental source: cv. Carina, leaf
C:Genetics:
A:Gene: CP29
C:Superfamily: chlorophyll a/b-binding protein

C:Keywords: chlorophyll; chloroplast; light-harvesting complex; photosystem II; thylakoid

Query Match 100.0%; Score 21; DB 2; Length 75;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
||||
DB 30 HARL 33

RESULT 8

F82684

hypothetical protein XF1417 [imported] - Xylella fastidiosa (strain 9a5c)

C:Species: Xylella fastidiosa

C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000

C:Accession: F82684

R:Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen

Nature 406, 151-157, 2000

A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.

A:Reference number: AB2515; PMID:20365717; PMID:10910347

A:Note: for a complete list of authors see reference number A59328 below

A:Accession: F82684

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-80 <SIM>

A:Cross-references: GB:AE003972; GB:AE003849; NID:g9106418; PIDN:AAF84226.1; GSPDB:GN001

A:Experimental source: strain 9a5c

R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A

Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carver, H

as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.

submitted to GenBank, June 2000

A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm

J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig

chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E

A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.

, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A

Rodrigues, V.; Rosa, A.J.; de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak

M.; Tsunako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z

A:Reference number: A59328

A:Contents: annotation

C:Genetics:

A:Gene: XF1417

Query Match

Best Local Similarity 100.0%; Score 21; DB 2; Length 80;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
||||
DB 53 HARL 56

RESULT 9

hypothetical protein BH3959 [imported] - Bacillus halodurans (strain C-125)

C:Species: Bacillus halodurans

C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001

C:Accession: G84144

R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hir

Nucleic Acids Res. 28, 4317-4331, 2000

A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and

A:Reference number: AB3650; PMID:120512582; PMID:11058132

A:Accession: G84144

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-83 <STO>

A:Cross-references: GB:AP001520; GB:BA000004; NID:g10176401; PIDN:BA07678.1; GSPDB:GN00

C:Genetics:

A:Gene: BH3959

Query Match 100.0%; Score 21; DB 2; Length 83;

Best Local Similarity 100.0%; Pred. No. 1.3e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
||||
DB 35 HARL 38

RESULT 10

H97404

hypothetical protein AGR_C_659 [imported] - Agrobacterium tumefaciens (strain C58, (

C:Species: Agrobacterium tumefaciens

C>Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 11-Jan-2002

C:Accession: H97404

R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quirillo, B.; Gol

A.; Liu, F.; Wollam, C.; Allinger, M.; Doughy, D.; Scott, C.; Lappas, C.; Markelz,

Science 294, 2323-2328, 2001

A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium

A:Reference number: A97359; PMID:11743194

A:Accession: H97404

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-85 <KUR>

A:Cross-references: GB:AE007869; PIDN:AAK86193.1; PID:g15155290; GSPDB:GN00169

C:Genetics:

A:Gene: AGR_C_659

A:Map position: circular chromosome

Query Match

Best Local Similarity 100.0%; Score 21; DB 2; Length 85;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
||||
DB 7 HARL 10

RESULT 11

AH2622

conserved hypothetical protein Atu0376 [imported] - Agrobacterium tumefaciens (stra

C:Species: Agrobacterium tumefaciens

C>Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002

C:Accession: AH2622

R:Wood, D.W.; Secubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; W

erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; Mc

Karp, P.; Romero, P.; Zhang, S.

Science 294, 2317-2323, 2001

A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-K

ster, E.W.

A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A:Reference number: AB2577; PMID:11743193

A:Accession: AH2622

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-85 <KUR>

A:Cross-references: GB:AE008688; PIDN:AAU41398.1; PID:g17738716; GSPDB:GN00186

A:Experimental source: strain C58 (Dupont)

C:Genetics:

A:Gene: Atu0376

A:Map position: circular chromosome

Query Match

Best Local Similarity 100.0%; Score 21; DB 2; Length 85;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
||||
DB 7 HARL 10

RESULT 12

EB3358

```
hypothetical protein PA2292 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: E83358
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Bhatman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Linn, J.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen
A:Reference number: AB2950; MUID:20437337; PMID:10984043
A:Accession: E83358
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-93 <STO>
A:Cross-references: GB:AE004655; GB:AE004091; NID:g9948321; PIDN:AA05680.1; GSPDB:GN001
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA2292

Query Match 100.0%; Score 21; DB 2; Length 93;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HARL 4
DB 31 HARL 34

RESULT 13
T20227
hypothetical protein C54G10.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T20227
R:Matthews, L.
submitted to the EMBL Data Library, June 1996
A:Reference number: Z19240
A:Accession: T20227
A:Status: preliminary; translated from GB/EMBL/DBJ
C:Accession: C81012; G81953
A:Molecule type: DNA
A:Residues: 1-93 <ML>
A:Cross-references: EMBL:Z75532; PIDN:CAA99809.1; GSPDB:GN00023; CESP:C54G10.1
A:Experimental source: clone C54G10
C:Genetics:
A:Gene: CESP:C54G10.1
A:Map position: 5

Query Match 100.0%; Score 21; DB 2; Length 93;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HARL 4
DB 19 HARL 22

RESULT 14
S38916
hypothetical protein 4 - phage phi-C31
C:Species: phage phi-C31
C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 08-Oct-1999
C:Accession: S38916
R:Hartley, N.M.; Murphy, G.O.; Bruton, C.J.; Chater, K.F.
submitted to the EMBL Data Library, November 1993
A:Reference number: S38912
A:Accession: S38916
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-96 <HAR>
A:Cross-references: EMBL:X76288; NID:g432610; PIDN:CAA53915.1; PID:g432615

Query Match 100.0%; Score 21; DB 2; Length 96;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HARL 4
DB 77 HARL 80

RESULT 16
C81012
conserved hypothetical protein NMB2058 [imported] - Neisseria meningitidis (strain MC
C:Species: Neisseria meningitidis
C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 02-Feb-2001
C:Accession: C81012; G81953
R:Reitelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, Hickey, E.K.; Haft, D.H.; Salberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.
rt, H.; Qin, H.; Yamathayan, J.; Gill, J.; Scariato, V.; Maignani, V.; Plaza, M.
Science 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.;
A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A:Reference number: AB1000; MUID:20175755; PMID:10710307
A:Accession: C81012
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-101 <TET>
A:Cross-references: GB:AE002555; GB:AE002098; NID:g7227310; PIDN:AAE42378.1; PID:g722
A:Experimental source: serogroup B, strain MC58
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Mc
; Holroyd, S.; Jorgels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandre
Nature 404, 502-506, 2000
A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491
A:Reference number: AB1775; MUID:20222556; PMID:10761919
A:Accession: G81953
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-101 <PAR>
A:Cross-references: GB:AL162753; GB:AL157959; NID:g7379120; PIDN:CAB83678.1; PID:g737
A:Experimental source: serogroup A, strain Z2491
C:Genetics:
A:Gene: NMB2058; MMA0377

Query Match 100.0%; Score 21; DB 2; Length 101;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HARL 4
DB 77 HARL 80

RESULT 15
A97833
hypothetical protein RC1065 [imported] - Rickettsia conorii (strain Malish 7)
C:Species: Rickettsia conorii
C>Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 30-Sep-2001
C:Accession: A97833
R:Ogata, H.; Audic, S.; Reneato-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.;
Science 293, 2093-2098, 2001
A:Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.
A:Reference number: A97700; MUID:21442074; PMID:11557893
A:Accession: A97833
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-100 <RUR>
A:Cross-references: GB:AE006914; PIDN:AL03603.1; PID:g15620185; GSPDB:GN00173
C:Genetics:
A:Gene: RC1065

Query Match 100.0%; Score 21; DB 2; Length 100;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HARL 4
DB 77 HARL 80
```

Db 12 HARL 15

RESULT 17

A97484

hypothetical protein AGR_C.1876 [imported] - Agrobacterium tumefaciens (strain C58, Cere
C:Species: Agrobacterium tumefaciens
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 11-Jan-2002

C:Accession: A97484
R:Gouder, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,

A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001

A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A:Reference number: A97359; PMID:11743194

A:Accession: A97484

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-103 <KOR>

A:Cross-references: GB:AE007869; PIDN:AAK8626.1; PID:g15156036; GSPDB:GN00169

C:Genetics:

A:Gene: AGR_C.1876

A:Map position: circular chromosome

Query Match

Best Local Similarity 100.0%; Score 21; DB 2; Length 103;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4

Db 61 HARL 64

RESULT 18

A12701

hypothetical protein Atu1018 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)

C:Species: Agrobacterium tumefaciens

C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002

C:Accession: A12701

R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I

Erange, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavlin, T.; Levy, R.; Li, M.; McClell

; Karp, P.; Romero, P.; Zhang, S.

Science 294, 2317-2323, 2001

A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,

ster, E.W.

A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A:Reference number: AB2577; PMID:11743193

A:Accession: A12701

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-103 <KOR>

A:Cross-references: GB:AE008688; PIDN:AAL42031.1; PID:g17739407; GSPDB:GN00186

A:Experimental source: strain C58 (Dupont)

C:Genetics:

A:Gene: Atu1018

A:Map position: circular chromosome

Query Match

Best Local Similarity 100.0%; Score 21; DB 2; Length 103;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4

Db 61 HARL 64

RESULT 19

D97412

hypothetical protein AGR_C.774 [imported] - Agrobacterium tumefaciens (strain C58, Cere
C:Species: Agrobacterium tumefaciens
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 11-Jan-2002

C:Accession: D97412

R:Gouder, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,

A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.

Science 294, 2323-2328, 2001

A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium

A:Reference number: A97359; PMID:11743194

A:Accession: D97412

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-103 <KOR>

A:Cross-references: GB:AE007869; PIDN:AAK86253.1; PID:g15155361; GSPDB:GN00169

C:Genetics:

A:Gene: AGR_C.774

A:Map position: circular chromosome

Query Match

Best Local Similarity 100.0%; Score 21; DB 2; Length 103;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4

Db 45 HARL 48

RESULT 20

S04786

mobilization protein mbc - Escherichia coli plasmid ColK

C:Species: Escherichia coli

C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 18-Jun-1993

C:Accession: S04786

R:Boyd, A.C.; Archer, J.A.K.; Sherratt, D.J.

Mol. Gen. Genet. 217, 488-498, 1989

A:Title: Characterization of the ColE1 mobilization region and its protein products.

A:Reference number: J00389; MUID:89364735; PMID:2671664

A:Accession: S04786

A:Status: not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-106 <BOY>

C:Genetics:

A:Gene: mbc; mbx

A:Genome: plasmid

Query Match

Best Local Similarity 100.0%; Score 21; DB 2; Length 106;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4

Db 11 HARL 14

RESULT 21

J00389

mobilization protein mbec - Escherichia coli plasmid ColE1

C:Species: Escherichia coli

C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 08-Oct-1999

C:Accession: J00389

R:Boyd, A.C.; Archer, J.A.K.; Sherratt, D.J.

Mol. Gen. Genet. 217, 488-498, 1989

A:Title: Characterization of the ColE1 mobilization region and its protein products.

A:Reference number: J00389; MUID:89364735; PMID:2671664

A:Accession: J00389

A:Molecule type: DNA

A:Residues: 1-107 <BOY>

A:Cross-references: GB:X15873; NID:945448; PIDN:CAA3382.1; PID:g45449

C:Genetics:

A:Gene: mbec

A:Genome: plasmid

Query Match

Best Local Similarity 100.0%; Score 21; DB 2; Length 107;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4

Db 11 HARL 14

RESULT 22

T07540
hypothetical protein 107 - Japanese black pine chloroplast

C:Species: chloroplast pins thunbergiana (japanese black pine)
C:Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 18-Aug-2000
C:Accession: T07540

R:Nakazuki, T.; Tsudzuki, J.; Ito, S.; Nakashima, K.; Tsudzuki, T.; Sugitara, M.
Proc. Natl. Acad. Sci. U.S.A. 91, 9794-9798, 1994

A:Title: Loss of all ndh genes as determined by sequencing the entire chloroplast genome

A:Reference number: Z16030; MUID:95024047; PMID:7937893

A:Accession: T07540

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-107 <NA>

A:Cross-references: EMBL:DI7510; NID:9529643; PIDN:BAA04416.1; PID:91262701

C:Genetics:

A:Genome: chloroplast

C:Keywords: chloroplast

Query Match

Best Local Similarity 100.0%; Score 21; DB 2; Length 107;
Pred. No. 1.7e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4

DB 75 HARL 78

RESULT 23

S43154
diaminopimelate epimerase (EC 5.1.1.7) - Pseudomonas aeruginosa (fragment)

C:Species: Pseudomonas aeruginosa

C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 16-Jul-1999

C:Accession: S61400; S43154

R:Hoefle, M.; Dong, O.; Kourambas, S.; Krishnapillai, V.; Sherratt, D.; Mergey, M.

Mol. Microbiol. 14, 1011-1020, 1994

A:Title: The ssb gene product, which affects pyoverdinin production in Pseudomonas aeruginosa

A:Reference number: S61400; MUID:95231283; PMID:7715441

A:Accession: S61400

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-108 <HO>

A:Cross-references: EMBL:X78478; NID:9468712; PIDN:CAA55224.1; PID:9468713

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, March 1994

R:Hoefle, M.; Dong, O.; Kourambas, S.; Krishnapillai, V.; Mergey, M.

submitted to the EMBL Data Library, March 1994

A:Description: The ssb gene product which affects pyoverdinin production in Pseudomonas aeruginosa

A:Reference number: S43154

A:Accession: S43154

A:Molecule type: DNA

A:Residues: 1-108 <HO>

A:Cross-references: EMBL:X78478; NID:9468712; PIDN:CAA55224.1; PID:9468713

C:Genetics:

A:Superfamily: diaminopimelate epimerase

C:Keywords: isomerase; lysine biosynthesis

Query Match

Best Local Similarity 100.0%; Score 21; DB 2; Length 108;
Pred. No. 1.7e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4

DB 35 HARL 38

RESULT 24

Q0CVB1

V2 protein - tomato yellow leaf curl virus

C:Species: tomato yellow leaf curl virus

C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 23-Jul-1999

C:Accession: E40779

R:Navot, N.; Pichersky, E.; Zeidan, M.; Zamir, D.; Czosnek, H.

Virology 185, 151-161, 1991

A:Title: tomato yellow leaf curl virus: a whitefly-transmitted geminivirus with a single

A:Reference number: A40779; MUID:92024070; PMID:1926771

A:Accession: E40779

A:Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-116 <NA>

A:Cross-references: GB:X15656; NID:962204; PIDN:CAA33687.1; PID:962206

C:Superfamily: tomato yellow leaf curl virus V2 protein

Query Match 100.0%; Score 21; DB 1; Length 116;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4

DB 65 HARL 68

RESULT 25

D49339
cell growth regulatory protein ChpB - Escherichia coli (strain K-12)

N:Alternate names: pemk-like protein 2

C:Species: Escherichia coli

C:Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 01-Mar-2002

C:Accession: D49339; S56451; D65234

R:Masuda, Y.; Miyakawa, K.; Nishimura, Y.; Ohtsubo, E.

J. Bacteriol. 175, 6850-6856, 1993

A:Title: chpA and chpB, Escherichia coli chromosomal homologs of the pem locus respon

A:Reference number: A49339; MUID:94042847; PMID:8226627

A:Accession: D49339

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-116 <NA>

A:Cross-references: GB:D16451; NID:9452513; PIDN:BAA03920.1; PID:d1004436; PID:950611

R:Burland, V.; Plunkett III, G.; Sofia, H.J.; Daniels, D.L.; Blattner, F.R.

Nucleic Acids Res. 23, 2105-2119, 1995

A:Title: Analysis of the Escherichia coli genome VI: DNA sequence of the region from

A:Reference number: S56314; MUID:95334362; PMID:7610040

A:Accession: S56451

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-116 <NA>

A:Cross-references: EMBL:U14003; NID:91263172; PIDN:AAA97122.1; PID:9537067

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1994

R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.

A:Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of Escherichia coli K-12.

A:Reference number: A64720; MUID:97426617; PMID:9278503

A:Accession: D65234

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-116 <BLAT>

A:Cross-references: GB:AE000494; GB:U00096; NID:91790670; PIDN:AACT7182.1; PID:917906

A:Experimental source: strain K-12, substrain MG1655

C:Genetics:

A:Gene: chpB

C:Keywords: DNA binding

Query Match 100.0%; Score 21; DB 2; Length 116;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4

DB 89 HARL 92

RESULT 26

I57506

oncogene tyrosine protein kinase receptor - human (fragment)

C:Species: Homo sapiens (man)
C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 05-Nov-1999
C:Accession: 157506
R:Collier, F.; Kumar, R.; Ernst, M.; Klein, R.; Martin-Zanca, D.; Barbacid, M.
Mol. Cell. Biol. 10, 4202-4210, 1990
A:Title: Human trk oncogenes activated by point mutation, in-frame deletion, and duplication
A:Reference number: 157506; MUID:90318386; PMID:1695324
A:Accession: 157506
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-118 <RES>
A:Cross-references: GB:M55336; NID:g3339913; PIDN:AAA36768.1; PID:g3339914
C:Genetics:
A:Gene: trk2

Query Match 100.0%; Score 21; DB 2; Length 118;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
||||
DB 7 HARL 10

RESULT 27

177386
oncogene tyrosine protein kinase receptor - human (fragment)

C:Species: Homo sapiens (man)
C>Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 05-Nov-1999
C:Accession: 177386
R:Collier, F.; Kumar, R.; Ernst, M.; Klein, R.; Martin-Zanca, D.; Barbacid, M.
Mol. Cell. Biol. 10, 4202-4210, 1990
A:Title: Human trk oncogenes activated by point mutation, in-frame deletion, and duplication
A:Reference number: 157506; MUID:90318386; PMID:1695324
A:Accession: 177386
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-118 <RES>
A:Cross-references: GB:M55337; NID:g3339915; PIDN:AAA36769.1; PID:g3339916
C:Genetics:
A:Gene: trk4

Query Match 100.0%; Score 21; DB 2; Length 118;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
||||
DB 7 HARL 10

RESULT 28

E70718
hypothetical protein RV0967 - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis
C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C:Accession: E70718
R:Cole, S.T.; Broch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltham, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: E70718
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-119 <COL>
A:Cross-references: GB:Z79700; GB:AL123456; NID:g3261628; PIDN:CAB02009.1; PID:e1300050;
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: RV0967

Query Match 100.0%; Score 21; DB 2; Length 119;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
||||
DB 96 HARL 99

RESULT 29

C96008
hypothetical membrane protein [imported] - Sinorhizobium meliloti (strain 1021) magc

C:Species: Sinorhizobium meliloti
C>Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
C:Accession: C96008
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; He
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A:Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing
A:Reference number: A95842; MUID:21396508; PMID:11481431
A:Accession: C96008
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-120 <KUP>
A:Cross-references: GB:AL59185; PIDN:CAC49731.1; PID:g15141218; GSPDB:GND0167
A:Experimental source: strain 1021, megaplasmid pSymB
R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hu
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, F
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lele
habault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Welis, D.H.; Wong, K.; Yel
A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A:Reference number: A96039; MUID:21368234; PMID:11474104
A:Contents: annotation
C:Genetics:
A:Gene: SMD21501
A:Genome: plasmid

Query Match 100.0%; Score 21; DB 2; Length 120;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
||||
DB 84 HARL 87

RESULT 30

E87139
50S ribosomal protein L18 [imported] - Mycobacterium leprae

C:Species: Mycobacterium leprae
C>Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 10-May-2001
C:Accession: E87139
R:Cole, S.T.; Elgimeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.
R.; Davies, R.M.; Devlin, K.; Duthey, S.; Feltham, T.; Fraser, A.; Hamlin, N.; Hol
eam, M.A.; Rutherford, K.M.
Nature 409, 1007-1011, 2001
A:Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.
A:Title: Massive gene decay in the leprosy bacillus.
A:Reference number: A86909; MUID:21128732; PMID:11234002
A:Accession: E87139
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-121 <STO>
A:Cross-references: GB:AL450380; NID:g13093539; PIDN:CAC30797.1; GSPDB:GND0147
C:Genetics:
A:Gene: rplR
C:Superfamily: Escherichia coli ribosomal protein L18

Query Match 100.0%; Score 21; DB 2; Length 121;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
 DB 16 HARL 19

RESULT 31

T37737

hypothetical protein SPAC167.06c - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jan-2000

C:Accession: T37737, T38952

R:Rieger, M.; Wood, V.; Rajandream, M.A.; Barrell, B.G.

submitted to the EMBL Data Library, January 1999

A:Reference number: Z21743

A:Accession: T37737

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-121 <REF>

A:Cross-references: EMBL:AL035248; PIDN:CAA22851.1; GSPDB:GN00066; SPDB:SPAC167.06c

A:Experimental source: strain 972h; cosmid c167

R:Skellon, J.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V.

submitted to the EMBL Data Library, May 1997

A:Reference number: Z21819

A:Accession: T38952

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-121 <SKB>

A:Cross-references: EMBL:Z95396; PIDN:CAB08760.1; GSPDB:GN00066; SPDB:SPAC57A7.02c

A:Experimental source: strain 972h; cosmid c57A7

C:Genetics:

A:Gene: SPAC57A7.02c; SPAC167.06c

A:Map position: 1

Query Match

100.0%; Score 21; DB 2; Length 121;

Best Local Similarity 100.0%; Pred. No. 1; 9e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4

DB 35 HARL 38

RESULT 32

T45382

ribosomal protein L18 [imported] - Mycobacterium leprae

C:Species: Mycobacterium leprae

C:Date: 31-Jan-2000 #sequence_revision 31-Jan-2000 #text_change 18-Feb-2000

C:Accession: T45382

R:Coile, S.T.; Flesselles, B.; Honore, N.

submitted to the EMBL Data Library, August 1997

A:Reference number: Z22966

A:Accession: T45382

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-122 <COL>

A:Cross-references: EMBL:Z98756; PIDN:CAB11452.1

A:Experimental source: cosmid B2492

C:Genetics:

A:Note: rplR

C:Superfamily: Escherichia coli ribosomal protein L18

Query Match

100.0%; Score 21; DB 2; Length 122;

Best Local Similarity 100.0%; Pred. No. 2e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4

DB 17 HARL 20

RESULT 33

G81198

hypothetical protein NMB0438 [imported] - Neisseria meningitidis (strain MC58 serogr

C:Species: Neisseria meningitidis

C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001

C:Accession: G81198

R:Rettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen,

Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.

ri, H.; Olo, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizsa, M.

Science 287, 1809-1815, 2000

A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappunli, R.

A>Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.

A:Reference number: A81000; MUID:2015755; PMID:10710307

A:Accession: G81198

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-122 <TEP>

A:Cross-references: GB:AE002400; GB:AE002098; NID:97225659; PIDN:AAE40876.1; PID:9722.

A:Experimental source: serogroup B, strain MC58

C:Genetics:

A:Gene: NMB0438

Query Match

100.0%; Score 21; DB 2; Length 122;

Best Local Similarity 100.0%; Pred. No. 2e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4

DB 76 HARL 79

RESULT 34

C81775

hypothetical protein NMA2047 [imported] - Neisseria meningitidis (strain Z2491 serogr

C:Species: Neisseria meningitidis

C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001

C:Accession: C81775

R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Mc

R:Holroyd, S.; Jagsels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandre

Nature 404, 502-506, 2000

A>Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491

A:Reference number: A81775; MUID:2022556; PMID:10761919

A:Accession: C81775

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-122 <PAR>

A:Cross-references: GB:AL162758; GB:AL157959; NID:97380672; PIDN:CAB85265.1; PID:9738

A:Experimental source: serogroup A, strain Z2491

C:Genetics:

A:Gene: NMA2047

Query Match

100.0%; Score 21; DB 2; Length 122;

Best Local Similarity 100.0%; Pred. No. 2e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4

DB 76 HARL 79

RESULT 35

D85843

unknown protein encoded within prophage CP-933V [imported] - Escherichia coli (strai

C:Species: Escherichia coli

C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001

C:Accession: D85843

R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May

iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dialmalanta, E.; Potamousis, K.; Apod

Nature 409, 529-533, 2001

A>Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A:Reference number: A85480; MUID:21074935; PMID:11206551

A:Accession: D85843

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-124 <STO>

A:Cross-references: GB:BA0005174; NID:g12516372; PIDN:AA657208.1; GSPDB:GN00145; UWGP:233
A:Experimental source: strain 0157:H7, substrain EDJ933
C:Genetics:
A:Gene: Z3320

Query Match 100.0%; Score 21; DB 2; Length 124;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
Db 104 HARL 107

RESULT 36
D85630
Probable tail assembly chaperone [imported] - Escherichia coli (strain 0157:H7, substrain
C:Species: Escherichia coli
C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 27-Nov-2001
C:Accession: D85630
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shingawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli 0157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: D85630
A:Molecule type: DNA
A:Status: preliminary
A:Residues: 1-124 <STO>
A:Cross-references: GB:AE005174; NID:g12514215; PIDN:AA655504.1; GSPDB:GN00145; UWGP:213
A:Experimental source: strain 0157:H7, substrain EDJ933
C:Genetics:
A:Gene: Z1371

Query Match 100.0%; Score 21; DB 2; Length 124;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
Db 104 HARL 107

RESULT 37
F90822
Probable tail assembly chaperone [imported] - Escherichia coli (strain 0157:H7, substrain
C:Species: Escherichia coli
C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
C:Accession: F90822
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shingawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli 0157:H7 and gene
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: F90822
A:Molecule type: DNA
A:Status: preliminary
A:Residues: 1-124 <HAY>
A:Cross-references: GB:BA000007; PIDN:BA034973.1; PID:g13361014; GSPDB:GN00154
A:Experimental source: strain 0157:H7, substrain RMD 0509952
C:Genetics:
A:Gene: EC61550

Query Match 100.0%; Score 21; DB 2; Length 124;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
Db 104 HARL 107

RESULT 38
Query Match 100.0%; Score 21; DB 2; Length 124;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

A90854
Probable tail assembly chaperone [imported] - Escherichia coli (strain 0157:H7, sub:
C:Species: Escherichia coli
C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
C:Accession: A90854
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han,
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shingawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli 0157:H7 and
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: A90854
A:Molecule type: DNA
A:Status: preliminary
A:Residues: 1-124 <HAY>
A:Cross-references: GB:BA000007; PIDN:BA035224.1; PID:g13361266; GSPDB:GN00154
A:Experimental source: strain 0157:H7, substrain RMD 0509952
C:Genetics:
A:Gene: EC61801

Query Match 100.0%; Score 21; DB 2; Length 124;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
Db 104 HARL 107

RESULT 39
B90909
Probable tail assembly chaperone [imported] - Escherichia coli (strain 0157:H7, sub:
C:Species: Escherichia coli
C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
C:Accession: B90909
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han,
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shingawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli 0157:H7 and
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: B90909
A:Molecule type: DNA
A:Status: preliminary
A:Residues: 1-124 <HAY>
A:Cross-references: GB:BA000007; PIDN:BA035665.1; PID:g13361708; GSPDB:GN00154
A:Experimental source: strain 0157:H7, substrain RMD 0509952
C:Genetics:
A:Gene: EC62242

Query Match 100.0%; Score 21; DB 2; Length 124;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
Db 104 HARL 107

RESULT 40
E90876
Probable tail assembly chaperone [imported] - Escherichia coli (strain 0157:H7, sub:
C:Species: Escherichia coli
C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
C:Accession: E90876
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han,
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shingawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli 0157:H7 and
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: E90876
A:Molecule type: DNA
A:Status: preliminary
A:Residues: 1-124 <HAY>
A:Cross-references: GB:BA000007; PIDN:BA035404.1; PID:g13361446; GSPDB:GN00154

Query Match 100.0%; Score 21; DB 2; Length 124;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
Db 104 HARL 107

RESULT 40
E90876
Probable tail assembly chaperone [imported] - Escherichia coli (strain 0157:H7, sub:
C:Species: Escherichia coli
C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
C:Accession: E90876
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han,
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shingawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli 0157:H7 and
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: E90876
A:Molecule type: DNA
A:Status: preliminary
A:Residues: 1-124 <HAY>
A:Cross-references: GB:BA000007; PIDN:BA035404.1; PID:g13361446; GSPDB:GN00154

A; Experimental source: strain O157:H7, substrain RIMD 0509952
C; Genetics:
A; Gene: ECs1981

Query Match	100.0%;	Score 21;	DB 2;	Length 124;
Best Local Similarity	100.0%;	Pred. No. 2e+02;		
Matches	4;	Conservative	0;	Mismatches
			0;	Indels
				Gaps
				0;

QY	1	HARL	4
Db	104	HARL	107

RESULT 41
T49512
hypothetical protein B14D6.640 (Imported) - *Neurospora crassa*
C1:Species: *Neurospora crassa*
03-11-2020 14:02:03 Tue Nov 03 2020 14:02:03

C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 18-Aug-2000
C:Accession: T49512
R:Schulte, U.; Aign, V.; Hohelsel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura
submitted to the Protein Sequence Database, May 2000
A:Reference number: Z25022
Accession number: Z25022

C: Superfamily: Neurospora crassa hypothetical protein BL4D6.640

Query Match	100.0%	Score 21;	DB 2;	Length 125;
Best Local Similarity	100.0%	Pred. No. 2e+02;		
Matches	4;	Conservative 0;	Mismatches 0;	Indels 0;
			Gaps 0;	

QY	1	HARL	4
Db	11	HARL	14

RESULT 42
PC2265
cytochrome P450 protein, CYP4H5 - Anopheles albimanus (fragment)

C:Date: 24-Feb-1995 #sequence_revision 24-Feb-1995 #text_change 11-Jun-1999
C:Accession: PC2265
R:Scott, J.A.; Collins, F.H.; Feyereisen, R.
Biochem. Biophys. Res. Commun. 205, 1452-1459, 1994
A:Title: Diversity of cytochrome P450 genes in the mosquito, *Anopheles albimanus*
A:Reference number: PC2257; M01D:95100961; PMID:7545368

Query Match	100.0%	Score 21	DB 2	Length 127
Best Local Similarity	100.0%	Pred. No.	2e+02	
Matches 4	Conservative 0	Mismatches 0	Indels 0	Gaps 0

QY	1	HART	4
Db	41	HART	44

```
RESULT 43
B91216
thioredoxin 1 [imported] - Escherichia coli (strain O157:H7, substrain RMD 0509952)
```

C:Species: Escherichia coli
C:date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
C:Accession: B91218

R.Hayashi, T. Makino, K. Ohnishi, M. Kurokawa, K. Ishii, K. Yokoyama, K. Han, G. Asahara, N. Yasunaga, T. Kubara, S. Shiba, T. Hattori, M. Shinagawa, H. DNA Res. 8, 11-22, 2001
 A1Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and A1Reference number: A59629; MUID:21156231; PMID:11258796

Query Match	100.0%	Score 21	DB 2	Length 127
Best Local Similarity	100.0%	Score NC	2e+02	
Matches 4	Conservative 0	Mismatches 0	Indels 0	Gaps 0

QY	1 HARL 4
Db	9 HARL 12

```
RESULT 44
CB6064
thioredoxin_1 [imported] - Escherichia coli (strain O157:H7,
substrain EDL933)
```

C.Species: *Escherichia coli*
C.Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C.Accession: C86064
R.Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Ma
liller, L.; Grobeck, E.J.; Davis, N.W.; Lam, A.; Diallanla, E.; Potlousists, K.; Apod
Nature 409, 529-533, 2001
A.Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.
A.Reference number: A85480; MUID:21074935; PMID:11206551

Query Match	100.0%	Score 21	DB 2	Length 127
Best Local Similarity	100.0%	Pred. NC	2e+02	
Matches 4	Conservative 0	Mismatches 0	Indels 0	Gaps 0

QY	1 HARL 4
Db	9 HARL 12

RESULT 45
G72460
hypothetical protein APE2331 - Aeropyrum pernix (strain K1)

C:Species: Aeropyrum pernix
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999
C:Accession: G72460
R:KikukawaYasui, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Halkawa, Y.; Jin-no, K.; T-
dawa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.
DNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aer
A:Reference number: A72450, MUID:99310339, PMID:10382966

A:Experimental source: strain K1
C:Genetics:

A:Gene: APE2331

Query Match

Best Local Similarity 100.0%; Score 21; DB 2; Length 129;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
||||

DB 44 HARL 47

RESULT 46

hypoetical protein D-132 - Escherichia coli plasmid pBR322

C:Species: Escherichia coli

C>Date: 31-Dec-1980 #sequence_revision 31-Dec-1980 #text_change 10-Sep-1999

C:Accession: A04481

R: Sutcliffe, J.G.

Cold Spring Harb. Symp. Quant. Biol. 43, 77-90, 1979

A:Title: Complete nucleotide sequence of the Escherichia coli plasmid pBR322.

A:Reference number: A90923; MID:80002802; PMID:383387

A:Accession: A04481

A:Molecule type: DNA

A:Residues: 1-132 <STO>

C:Genetics:

A:Genome: plasmid

C:Superfamily: Escherichia coli plasmid pBR322 hypothetical 15.3k protein

Query Match

Best Local Similarity 100.0%; Score 21; DB 1; Length 132;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
||||

DB 32 HARL 35

RESULT 47

hypoetical protein At2g27180 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001

C:Accession: G84669

R: Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.

euens, D.; Nierman, M.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MID:20083487; PMID:10617197

A:Accession: G84669

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-132 <STO>

A:Cross-references: GB:AE002093; MID:94646220; PIDN:AAD26886.1; GSPDB:GN00139

C:Genetics:

A:Gene: At2g27180

A:Map position: 2

Query Match

Best Local Similarity 100.0%; Score 21; DB 2; Length 132;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
||||

DB 75 HARL 78

RESULT 48

T07498

hypothetical protein 133 - Japanese black pine chloroplast
C:Species: chloroplast Pinus thunbergiana (Japanese black pine)

C>Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 18-Aug-2000

C:Accession: T07498

R:Wakasugi, T.; Tsudzuki, J.; Ito, S.; Nakashima, K.; Tsudzuki, T.; Sugira, M.

Proc. Natl. Acad. Sci. U.S.A. 91, 9794-9798, 1994

A:Title: Loss of all ndh genes as determined by sequencing the entire chloroplast ge

A:Reference number: Z16030; MID:95024047; PMID:7937893

A:Accession: T07498

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-133 <MAX>

A:Cross-references: EMBL:D17510; MID:9529643; PIDN:BA04376.1; PID:g1262659

C:Genetics:

A:Genome: chloroplast

C:Superfamily: Japanese black pine chloroplast hypothetical protein 133

C:Keywords: chloroplast

Query Match

Best Local Similarity 100.0%; Score 21; DB 2; Length 133;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
||||

DB 94 HARL 97

RESULT 49

S72720

B1177-CL147 protein - Mycobacterium leprae

C:Species: Mycobacterium leprae

C>Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 23-Mar-2001

C:Accession: S72720

R:Smith, D.R.; Robison, K.

submitted to the EMBL Data Library, November 1993

A:Description: Mycobacterium leprae cosmid B1177.

A:Reference number: S72694

A:Accession: S72720

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-133 <SMI>

A:Cross-references: EMBL:U00011; MID:9466807; PIDN:AAAL7084.1; PID:9466809

C:Genetics:

A:Start codon: GTG

Query Match

Best Local Similarity 100.0%; Score 21; DB 2; Length 133;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
||||

DB 116 HARL 119

RESULT 50

B83475

hypothetical protein PA1353 [imported] - Pseudomonas aeruginosa (strain PA01)

C:Species: Pseudomonas aeruginosa

C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000

C:Accession: B83475

R:Stoyer, C.K.; Pham, X.O.; Eryin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.;

.; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic p

A:Reference number: A82950; MID:20437337; PMID:10984043

A:Accession: B83475

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-137 <STO>

A:Cross-references: GB:AE004565; GB:AE004091; MID:99947294; PIDN:AAG04742.1; GSPDB:G

A:Experimental source: strain PA01

C:Genetics:

A:Gene: PA1353

Query Match 100.0%; Score 21; DB 2; Length 137;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 HARL 4
DB 53 HARL 56

Search completed: November 13, 2002, 13:41:39
Job time : 17 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2002, 13:23:38 : Search time 10 Seconds
(without alignments)
16.591 Million cell updates/sec

Title: US-09-697-590-2_COPY_292_295
Perfect score: 21
Sequence: 1 HARL 4

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	21	100.0	71	Y698_ARCFU	O29560 archaeoglob
2	21	100.0	83	ES98_HUMAN	O9buw7 homo sapien
3	21	100.0	83	ES98_MOUSE	P58686 mus musculi
4	21	100.0	93	VP44_BPAPS	O9t1q4 bacterioph
5	21	100.0	107	MBEC_ECOLI	P13657 escherichia
6	21	100.0	116	CHPB_ECOLI	P33647 escherichia
7	21	100.0	116	YYL_TYICV	P27269 tomato yell
8	21	100.0	121	YWM2_SCHRO	P87133 schizosacch
9	21	100.0	122	RLH8_MYCLE	O32999 mycobacteri
10	21	100.0	132	YPI1_ECOLI	P03852 escherichia
11	21	100.0	133	Y145_ADE07	P05667 human adeno
12	21	100.0	141	NIFD_RHILT	P00465 rhizobium 1
13	21	100.0	155	RS19_DROME	P39018 drosophila
14	21	100.0	162	YATO_ECOLI	P75816 escherichia
15	21	100.0	163	PTPA_MYCTU	P75816 escherichia
16	21	100.0	169	RPC2_BP186	P21678 mycobacteri
17	21	100.0	182	RECA_ACEAC	O08784 bacterioph
18	21	100.0	197	DETD_CAEEL	P30648 caenorhabdi
19	21	100.0	197	RP34_BPSP1	P06227 bacterioph
20	21	100.0	207	COAE_XANCP	O56764 xanthomonas
21	21	100.0	208	COAE_XYLFA	O56764 xanthomonas
22	21	100.0	208	FUSD_BURCE	O9p412 xyella fas
23	21	100.0	208	KTHY_AERPE	O9p412 xyella fas
24	21	100.0	224	DKK4_HUMAN	O9p412 xyella fas
25	21	100.0	226	BIOD_XYLFA	O9p412 xyella fas
26	21	100.0	236	YP2C_STANU	P14503 staphylococ
27	21	100.0	241	HUTC_KLEAE	P14503 staphylococ
28	21	100.0	248	YR31_RALSO	O8yva0 ralsstonia s
29	21	100.0	256	CB4A_LYCES	P27524 lycopersico
30	21	100.0	258	KC2C_YEAST	P36930 saccharomyc
31	21	100.0	261	CB4_SPIOR	P36494 spinacia oi
32	21	100.0	266	NIFM_KLEPN	P08534 klebsiella
33	21	100.0	268	TP932_TREPA	O07950 treponema p

34	21	100.0	269	1	T2S1_STRPT	O52512 streptomyc
35	21	100.0	271	1	ICA4_HUMAN	O14773 homo sapien
36	21	100.0	273	1	NUPM_NEUCR	P24919 neuropept
37	21	100.0	274	1	RECA_NEUCR	O59595 neisseria c
38	21	100.0	274	1	YG73_SYNY3	P14261 synechocyst
39	21	100.0	275	1	CTA2_ACILM	O33950 actinobact
40	21	100.0	276	1	DAPF_PSEAE	O51564 pseudomonas
41	21	100.0	280	1	GSTG_BP12	O06717 bacterioph
42	21	100.0	280	1	GSTG_BP16	O06717 bacterioph
43	21	100.0	280	1	MDCB_PSEPU	O92452 pseudomonas
44	21	100.0	300	1	GP40_HUMAN	O14842 homo sapien
45	21	100.0	307	1	MMSR_PSEAE	P28805 pseudomonas

ALIGNMENTS

RESULT 1
Y698_ARCFU STANDARD: PRT: 71 AA.
ID Y698_ARCFU
AC O29560;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein AF0698.
GN AF0698.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
OC Archaeoglobaceae; Archaeoglobus.
OX NCBI_TaxID=2234;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kierlavage A.R., Graham D.E., Kyriades N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Oteybeck T.,
RA Cotton P.M., Spriggs T., Artlich P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
reducing archaeon Archaeoglobus fulgidus.";
RL Nature 390:364-370(1997).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
DR EMBL: AF001056; AAB90546.1; -
DR TIGR: AF0698;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 71 AA; 8002 MM; 803C030788157E24 CRC64;
Query Match 100.0%; Score 21; DB 1; Length 71;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HARL 4
DB 4 HARL 7

RESULT 2
ES98_HUMAN

```

ID ES98_HUMAN STANDARD; PRT; 83 AA.
AC Q9Y3F7.
DT 15-JUN-2002 (Rel. 41, Created)
DE 15-JUN-2002 (Rel. 41, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein EST00098.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=Muscle;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
RN (2)
RX MEDLINE=99299247; PubMed=10369878;
RA Gilley J., Fried M.;
RT "Extensive gene order differences within regions of conserved synteny
RT between the fugu and human genomes: implications for chromosomal
RL Hum. Mol. Genet. 8:1313-1320(1999).
CC -1- SIMILARITY: BELONGS TO THE UPF0184 (EST00098) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; BC001857; AA01857.1;
DR EMBL; BC008887; AA08887.1;
DR Gene; Y17450; CAB4344.1;
DR Gene; HGNC:17823; C9orf16.
DR InterPro; IPR005374; UPF0184.1.
DR Pfam; PF03670; UPF0184.1.
DR Hypothetical protein; Coiled coil.
KW DOMAIN 25 74 COILED COIL (POTENTIAL).
FT CONFLICT 35 35 L -> V (IN REF. 2).
SQ SEQUENCE 83 AA; 9054 MW; EBEID0AFAD6DADB CRC64;

Query Match 100.0%; Score 21; DB 1; Length 83;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HARL 4
DB 53 HARL 56

RESULT 3
ES98_MOUSE STANDARD; PRT; 83 AA.
ID ES98_MOUSE
AC P58686;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein EST00098.
OS Mus musculus (mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryo;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shuagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Iizawa M., Nishi K., Kiyosawa H., Kondo S., Yamanka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,

```

```

RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiya H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schirral L.M., Staubli F., Suzuki K., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Butt C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamita M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Momberts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohlsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC -1- SIMILARITY: BELONGS TO THE UPF0184 (EST00098) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AK003598; NOT_ANNOTATED_CDS.
DR MGI; MGI:1920987; 1110008P14RIK.
DR InterPro; IPR005374; UPF0184.
DR Pfam; PF03670; UPF0184.1.
DR Hypothetical protein; Coiled coil.
KW DOMAIN 25 74 COILED COIL (POTENTIAL).
FT CONFLICT 35 35 L -> V (IN REF. 2).
SQ SEQUENCE 83 AA; 9055 MW; 4D6D5BA98B72D36D CRC64;

Query Match 100.0%; Score 21; DB 1; Length 83;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HARL 4
DB 53 HARL 56

RESULT 4
VP44_BPAPS STANDARD; PRT; 93 AA.
ID VP44_BPAPS
AC O9Y104;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Putative protein P44.
DE 44.
GN Bacteriophage APSE-1.
OS Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;
OC unclassified Podoviridae.
OX NCBI_TaxID=106199;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=99420383; PubMed=10489345;
RA van der Wilk F., Dulleman A.M., Verbeek M., van den Heuvel J.F.J.M.;
RT "Isolation and characterization of APSE-1, a bacteriophage infecting
RT the secondary endosymbiont of acyrthosiphon pisum.";
RL Virology 262:104-113(1999).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF157835; AA03987.1;

```

KM Hypothetical protein.
SQ SEQUENCE 93 AA; 10356 MW; B7EEB57CB8BB3C63 CRC64;
Query Match 100.0%; Score 21; DB 1; Length 93;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 HARL 4
| | | |
DB 67 HARL 70
RESULT 5
ID MBEC_ECOLI STANDARD; PRT; 107 AA.
AC P13657;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 01-MAY-1992 (Rel. 22, Last annotation update)
DE Mobilization protein mbec.
GN MBEC.
OS Escherichia coli.
OG Plasmid COLE1.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89364735; PubMed=2671664;
RA Boyd A.C., Archer J.A.K., Sherratt D.J.;
RT "Characterization of the COLE1 mobilization region and its protein
RT products";
RL Mol. Gen. Genet. 217:488-498(1989).
CC -!- FUNCTION: THIS PROTEIN IS ESSENTIAL TO PROMOTE THE SPECIFIC
CC TRANSFER OF THE PLASMID IN THE PRESENCE OF CONJUGATIVE PLASMIDS.
CC -!- SIMILARITY: HIGH, TO THE MBAC AND MBEC PROTEINS OF E.COLI.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC DR EMBL; X15873; CAA33882.1; -
CC PIR; JQ0389; JQ0389.
KW Plasmid; Mobility protein; Conjugation.
SQ SEQUENCE 107 AA; 11856 MW; 1372246AD3472182 CRC64;
Query Match 100.0%; Score 21; DB 1; Length 107;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 HARL 4
| | | |
DB 11 HARL 14
RESULT 6
ID CHPB_ECOLI STANDARD; PRT; 116 AA.
AC P33647;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE PemK-like protein 2.
GN CHPB OR CHPBK OR B4225.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia
OX NCBI_TaxID=562;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN-K12;
RX MEDLINE=94042847; PubMed=8226627;
RA Masuda Y., Miyakawa K., Nishimura Y., Ohtsubo E.;
RT "chpA and chpB, Escherichia coli chromosomal homologs of the pem
RT locus responsible for stable maintenance of plasmid R100.";
RL J. Bacteriol. 175:6850-6856(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RX MEDLINE=95334362; PubMed=7610040;
RA Burdard V.D., Plunkett G. III, Sofia H.J., Daniels D.L.,
RA Blattner F.R.;
RT "Analysis of the Escherichia coli genome VI: DNA sequence of the
RT region from 92.8 through 100 minutes";
RL Nucleic Acids Res. 23:2105-2119(1995).
RN [3]
RP SEQUENCE OF 1-30 FROM N.A.
RC STRAIN-K12;
RX MEDLINE=86278038; PubMed=3525538;
RA Weiss D.L., Johnson D.I., Weith H.L., Somerville R.L.;
RT "Structural analysis of the 11er locus of Escherichia coli K12.";
RL J. Biol. Chem. 261:9966-9971(1986).
RN [4]
RP SEQUENCE OF 20-116 FROM N.A.
RC STRAIN-K12;
RX MEDLINE=89053923; PubMed=2848015;
RA Lahti R., Pitkaeranta T., Valve E., Ilta I., Kukko-Kalske E.,
RA Heino J.;
RT "Cloning and characterization of the gene encoding inorganic
RT pyrophosphatase of Escherichia coli K-12";
RL J. Bacteriol. 170:5901-5907(1988).
RN [5]
RP GENE MAPPING.
RX MEDLINE=94364970; PubMed=8083180;
RA Masuda Y., Ohtsubo E.;
RT "Mapping and disruption of the chpB locus in Escherichia coli.";
RL J. Bacteriol. 176:5861-5863(1994).
CC -!- FUNCTION: MAY BE INVOLVED IN THE REGULATION OF CELL GROWTH. IT ACT
CC AS A GROWTH INHIBITOR. BOTH CHPS AND CHPB BIND TO THE PROMOTER
CC REGION OF THE CHPB OPERON TO AUTOREGULATE THEIR SYNTHESIS.
CC -!- SIMILARITY: BELONGS TO THE PEMK FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC DR EMBL; D16451; BAA03920.1; -
CC EMBL; M14018; -; NOT_ANNOTATED_CDS.
DR EMBL; U14003; AAA97122.1; -
DR EMBL; AE000494; AAC77182.1; -
DR EMBL; M23550; -; NOT_ANNOTATED_CDS.
DR PIR; D49339; D49339.
DR EcoGene; EGI2096; chpB.
DR InterPro; IPR003477; PemK.
DR Pfam; PF02452; PemK; 1.
KW DNA-binding; Complete proteome.
SQ SEQUENCE 116 AA; 12492 MW; 4D641F3B302FCF58 CRC64;
Query Match 100.0%; Score 21; DB 1; Length 116;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 HARL 4
| | | |
DB 89 HARL 92
RESULT 7

```

YV1_TYICV
ID YV1_TYICV STANDARD: PRT: 116 AA.
AC P27269.
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE Hypothetical 13.5 kDa protein (Y1 protein).
GN Y1.
OS Tomato yellow leaf curl virus (TYLCV).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=10832;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92024070; PubMed=1926771;
RA Navot N., Pichersky E., Zeidan M., Zamir D., Cosmek H.;
RT "Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus
RL Virology 185:151-161(1991).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sb-sib.ch).
CC -----
DR EMBL; X15656; CAA33687.1; -.
DR PIR; E40779; QOCVBI.
DR InterPro: IPR002511; Geminl_V1.
DR InterPro: IPR005159; WCCH.
DR Pfam; PF01524; Geminl_V1; 1.
DR Pfam; PF03716; WCCH; 1.
DR ProDom; PD002978; Geminl_V1; 1.
DR Hypothetical protein.
KW SEQUENCE 116 AA; 13459 MW; D506FB53AAF534BB CRC64;

```

```

Query Match 100.0%; Score 21; DB 1; Length 116;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 HARL 4
    |||
Db 65 HARL 68

```

```

RESULT 8
YDM2_SCHPO STANDARD: PRT: 121 AA.
AC P87133.
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein C57A7.02c in chromosome I.
GN SPAC57A7.02C OR SPAC167.06C.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;

```

```

RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Squorois J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,

```

```

RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymoprez B.,
RA Wellens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wandt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaire V., Mottier S.,
RA Gilbert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado I., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerruti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakowski G.V., Usseery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe."
RL Nature 415:871-880(2002).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sb-sib.ch).
CC -----
DR EMBL; Z95396; CAB08760.1; -.
DR EMBL; AL035248; CAA22851.1; -.
KW Hypothetical protein.
SQ SEQUENCE 121 AA; 13357 MW; 0461FD5759AEB117 CRC64;

```

```

Query Match 100.0%; Score 21; DB 1; Length 121;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 HARL 4
    |||
Db 35 HARL 38

```

```

RESULT 9
RL18_MYCLE STANDARD: PRT: 122 AA.
AC O32939.
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 50S ribosomal protein L18.
GN RPL8 OR ML1843 OR MLCB2492.20.
OS Mycobacterium leprae.
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1769;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=TN;
RX MEDLINE=21128732; PubMed=11234002;

```

```

RA Cole S.T., Eiglmeyer K., Parkhill J., James K.D., Thomson N.R.,
RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R.M., Devlin K., Duthoy S., Fellwell T., Fraser A., Hamlin N.,
RA Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA Barrell B.G.;
RT "Massive gene decay in the leprosy bacillus."
RL Nature 409:1007-1011(2001).
CC -1- FUNCTION: THIS IS ONE OF 3 PROTEINS THAT MEDIANE THE ATTACHMENT OF
CC THE 5S RNA INTO THE LARGE RIBOSOMAL SUBUNIT (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE L18P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z98756; CAB1452.1; -
DR EMBL; AL583923; CAC30797.1; ALT_INIT.
DR Leproma; ML1843; -
DR InterPro; IPR004389; L18_bact.
DR InterPro; IPR001149; Ribosomal_L18p.
DR Pfam; PF00861; Ribosomal_L18p; 1.
DR ProDom; PD001394; Ribosomal_L18p; 1.
DR TIGFAMS; TIGR00060; L18_bact; 1.
DR Ribosomal protein; rRNA-binding; Complete proteome.
SQ SEQUENCE 122 AA; 13275 MW; FA3C344CA7063NAF CRC64;

Query Match
Best Local Similarity 100.0%; Score 21; DB 1; Length 122;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
   ||||
Db 17 HARL 20

RESULT 10
YPB1_ECOLI STANDARD; PRT; 132 AA.
AC P08532;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-NOV-1990 (Rel. 16, Last annotation update)
DE Hypothetical 15.3 kDa protein.
OS Escherichia coli.
OC Plasmid pBR322.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=80002802; PubMed=383387;
RA Sutcliffe J.G.;
RT "Complete nucleotide sequence of the Escherichia coli plasmid
RT pBR322."
RL Cold Spring Harb. Symp. Quant. Biol. 43:77-90(1979).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J01749; -; NOT_ANNOTATED_CDS.
DR PIR; A04481; QOEC8
KW Hypothetical protein; Plasmid.
SQ SEQUENCE 132 AA; 15322 MW; 03D4DA9D841B588C CRC64;

Query Match
Best Local Similarity 100.0%; Score 21; DB 1; Length 132;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
   ||||
Db 32 HARL 35

RESULT 11
Y145_ADE07 STANDARD; PRT; 133 AA.
ID Y145_ADE07

```

```

AC P05667;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 01-AUG-1992 (Rel. 23, Last annotation update)
DE Hypothetical 14.5 kDa early protein.
OS Human adenovirus type 7.
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX NCBI_TaxID=10519;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Gomen;
RX MEDLINE=83183660; PubMed=6301944;
RA Engler J.A., Hoppe W.S., van Bree M.P.;
RT "The nucleotide sequence of the genes encoded in early region 2b of
RT human adenovirus type 7."
RL Gene 21:145-159(1983).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X03000; CAA26771.1; -
DR Hypothetical protein; Early protein.
KW Hypothetical protein; Early protein.
SQ SEQUENCE 133 AA; 14557 MW; 3585A331C09060D1 CRC64;

Query Match
Best Local Similarity 100.0%; Score 21; DB 1; Length 133;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
   ||||
Db 95 HARL 98

RESULT 12
NIFD_RHILF STANDARD; PRT; 141 AA.
ID NIFD_RHILF
AC P00465;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Nitrogenase molybdenum-iron protein alpha chain (EC 1.18.6.1)
DE (Nitrogenase component 1) (Dinitrogenase) (Fragment).
GN NIFD.
OS Rhizobium leguminosarum (biovar trifolii).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=386;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SU329;
RX MEDLINE=83261209; PubMed=6307623;
RA Scott K.F., Rolfe B.G., Shine J.;
RT "Biological nitrogen fixation: primary structure of the Rhizobium
RT trifolii iron protein gene."
RL DNA 2:149-155(1983).
CC -----
CC -1- FUNCTION: THE KEY ENZYMOLOGIC REACTIONS IN NITROGEN FIXATION ARE
CC CATALYZED BY THE NITROGENASE COMPLEX, WHICH HAS 2 COMPONENTS: THE
CC IRON PROTEIN AND THE MOLYBDENUM-IRON PROTEIN.
CC -1- CATALYTIC ACTIVITY: 8 reduced ferredoxin + 8 H(+) + N(2) + 16 ATP
CC = 8 oxidized ferredoxin + 2 NH(3) + 16 ADP + 16 phosphate.
CC -1- SUBUNIT: Tetramer of two alpha and two beta chains that binds
CC 30-32 Fe, 2 Mo, and inorganic sulfur.
CC -1- SIMILARITY: BELONGS TO THE NIFD/NIFE/NIF/NIFN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; K00490; -; NOT_ANNOTATED_CDS.
 DR PIR; A00543; NIZRAT.
 DR HSSP; P07328; 3MIN.
 DR InterPro: IPR000318; Nitrogenase_comp1.
 DR InterPro: IPR000510; Oxired_nitrogensel.
 DR Pfam; PF00148; oxidored_nitro; 1.
 DR PROSITE; PS00090; NITROGENASE_1_2; PARTIAL.
 DR PROSITE; PS00699; NITROGENASE_1_1; 1.
 KM Oxidoreductase; Nitrogen fixation; Molybdenum; Iron-sulfur.
 FT NON_PRR 141 141
 SQ SEQUENCE 141 AA; 15524 MW; A89CE780C577A40E CRC64;

Query Match 100.0%; Score 21; DB 1; Length 141;
 Best Local Similarity 100.0%; Pred. No. 92;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
 ||||
 DB 12 HARL 15

RESULT 13
 RS19_DROME STANDARD; PRT; 155 AA.
 ID RS19_DROME
 AC P39018;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 01-FEB-1995 (Rel. 31, Last annotation update)
 DE 40S ribosomal protein S19.
 GN RPS19.
 OS Drosophila melanogaster (Fruit fly).
 OC Insecta; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 CC Eukaryota; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 NC NCBI_TaxID=1227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Canlon-S;
 RX MEDLINE=93376521; PubMed=8367309;
 RA Baumgartner S.W., Martin D., Chiquet-Ehrismann R.;
 RL Nucleic Acids Res. 21:3897-3897(1993).
 CC -1- SIMILARITY: BELONGS TO THE S19E FAMILY OF RIBOSOMAL PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; X73153; CAAS1677.1; -;
 DR FlyBase; FBgn0010412; RPS19.
 DR InterPro: IPR001266; Ribosomal_S19E.
 DR Pfam; PF01090; Ribosomal_S19e; 1.
 DR PRODOM; PD003854; Ribosomal_S19E; 1.
 DR PROSITE; PS00628; RIBOSOMAL_S19E; 1.
 KM Ribosomal protein.
 FT INT_MET 0
 SQ SEQUENCE 155 AA; 17174 MW; F77AB02833E7797F CRC64;
 BY SIMILARITY.

Query Match 100.0%; Score 21; DB 1; Length 155;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
 ||||
 DB 108 HARL 111

RESULT 14

YB00_ECOLI STANDARD; PRT; 162 AA;
 ID YB00_ECOLI
 AC P75816;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein yb00.
 GN YB00 OR B0858 OR Y1085 OR ECS0938.
 OS Escherichia coli, and
 OS Escherichia coli O157:H7.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 NC NCBI_TaxID=562, 83334;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12."
 RL Science 277:1453-1474(1997).
 [2]
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=97061202; PubMed=8905232;
 RA Oshima T., Alba H., Baba T., Fujita K., Hayashi K., Honjo A.,
 RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
 RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
 RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,
 RA Sempel G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
 RA Yano M., Horiiuchi T.;
 RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome
 RT corresponding to the 12.7-28.0 min region on the linkage map."
 RL DNA Res. 3:137-155(1996).
 [3]
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
 RX MEDLINE=21074935; PubMed=11206551;
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
 RA Postel G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
 RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Blattner F.R.;
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7."
 RL Nature 409:529-533(2001).
 [4]
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / RIMD 0509952;
 RX MEDLINE=21156231; PubMed=11258796;
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tode T.,
 RA Iida T., Takami H., Honda T., Sasekawa C., Ogasawara N., Yasunaga T.,
 RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
 RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
 RT O157:H7 and genomic comparison with a laboratory strain K-12."
 RL DNA Res. 8:11-22(2001).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

EMBL; AE000187; AAC73945.1; -;

DR EMBL: D90723; BAA35569.1; -
 DR EMBL: D90724; BAA35572.1; -
 DR EMBL: AE005266; AAG55234.1; -
 DR EMBL: AP002553; BAB34361.1; -
 DR Ecogene: EG33684; ybjo.
 KW Hypothetical protein; Transmembrane; Complete proteome.
 FT TRANSMEM 24 44 POTENTIAL.
 FT TRANSMEM 67 87 POTENTIAL.
 FT TRANSMEM 95 115 POTENTIAL.
 SO SEQUENCE 162 AA; 18518 MW; C618B11524CB0DCB CRC64;

Query Match
 Best Local Similarity 100.0%; Score 21; DB 1; Length 162;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
 DB 16 HARL 19

RESULT 15
 PP2A_MYCTU STANDARD; PRT; 163 AA.
 AC Q10507;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Probable low molecular weight protein-tyrosine-phosphatase
 (EC 3.1.3.48) (Prtase).
 GN PP2A OR RV2234 OR MT2293 OR MTC427.15.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
 OC Actinomycetales; Corynebacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-H37RV;
 RX MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skellon S., Squares S., Squares R.,
 RA Sultson J.E., Taylor K., Whitehead S., Barrett B.G.;
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence."
 RL Nature 393:537-544(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CDC 1551 / Oshkosh;
 RA Felschmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., Debey R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
 RA Delcher A., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
 RA Bishai W.;
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains."
 RL Submitted (APR-2001) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: ACTS ON TYROSINE PHOSPHORYLATED PROTEINS. LOW-MW ARYL
 CC PHOSPHATES AND NATURAL AND SYNTHETIC ACYL PHOSPHATES.
 CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
 CC tyrosine + phosphate.
 CC -1- SIMILARITY: BELONGS TO THE LOW MOLECULAR WEIGHT PHOSPHOTYROSINE
 CC PROTEIN PHOSPHATASES FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)

CC or send an email to license@isb-sib.ch.
 CC -----
 DR EMBL: Z70692; CAA94656.1; -
 DR EMBL: AE007074; AAK46577.1; -
 DR HSSP: P11064; IDC9.
 DR TIGR: MT2293; -
 DR TubercuList; RV2234; -
 DR InterPro: IPR00106; Low_mwt_PTPase.
 DR Pfam; PF01451; LMPC; 1.
 DR SMART; SM00226; LMPC; 1.
 KW Hydrolase; Complete proteome.
 FT ACT_SITE 11 11 NUCLEOPHILE (BY SIMILARITY).
 FT ACT_SITE 17 17 BY SIMILARITY.
 FT ACT_SITE 126 126 PROTON DONOR (BY SIMILARITY).
 SO SEQUENCE 163 AA; 17892 MW; 9D1CF0DB24CA8E27 CRC64;

Query Match
 Best Local Similarity 100.0%; Score 21; DB 1; Length 163;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
 DB 93 HARL 96

RESULT 16
 RP2_BP186 STANDARD; PRT; 169 AA.
 ID RP2_BP186
 AC P21678;
 DT 01-MAY-1991 (Rel. 18, Created)
 DT 01-MAY-1991 (Rel. 18, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE Regulatory protein CII.
 GN CII OR CP76.
 OS Bacteriophage 186.
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;
 OC P2-like viruses.
 OX NCBI_TaxID=29252;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=89199647; PubMed=2704042;
 RX Richardson H., Puspurs A., Egan J.B.;
 RT "Control of gene expression in the P2-related temperate coliphage
 RT 186. VI. Sequence analysis of the early lytic region."
 RL J. Mol. Biol. 206:251-255(1989).
 RN [2]
 RP SEQUENCE OF 1-147 FROM N.A.
 RC STRAIN-186CITSP;
 RX MEDLINE=87112711; PubMed=3806670;
 RA Kallionis B., Dodd I.B., Egan J.B.;
 RT "Control of gene expression in the P2-related temperate coliphages.
 RT III. DNA sequence of the major control region of phage 186.";
 RL J. Mol. Biol. 191:199-209(1986).
 CC -1- FUNCTION: INVOLVED IN THE ESTABLISHMENT OF LYSOGENY.
 CC -1- SIMILARITY: 70% IDENTITY TO ORF OF E. COLI RETRON EC67.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)
 CC or send an email to license@isb-sib.ch.
 CC -----
 DR EMBL: X15001; CAA33106.1; -
 DR EMBL: U32222; AAC34178.1; -
 DR PIR: S03593; S03593.
 KW Early protein; Transcription regulation; Activator; DNA-binding.
 SO SEQUENCE 169 AA; 18692 MW; 9026ADBC0CA9D37C CRC64;

Query Match
 Best Local Similarity 100.0%; Score 21; DB 1; Length 169;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RT polymerase.";
RL J. Mol. Biol. 180:533-547(1984).
CC -I- FUNCTION: THE SIGMA FACTOR IS AN INITIATION FACTOR THAT PROMOTES
CC ATTACHMENT OF THE RNA POLYMERASE TO SPECIFIC INITIATION SITES AND
CC THEN IS RELEASED.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X01807; CAA25950.1; -
DR PIR; S08577; S08577.
KM Transcription regulation; Sigma factor; DNA-directed RNA polymerase;
KW DNA-binding.
SQ SEQUENCE 197 AA; 23679 MW; 26D76E7548BDF0D CRC64;

Query Match 100.0%; Score 21; DB 1; Length 197;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HARL 4
DB 171 HARL 174

RESULT 20
COAE_XANCP STANDARD; PRT; 207 AA.
AC Q56764;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Dephospho-CoA kinase (EC 2.7.1.24) (Dephosphocoenzyme A kinase).
GN COAE OR XCC3102.
OS Xanthomonas campestris (pv. campestris).
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
CC Xanthomonas
CX NCBI_TaxID=340;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33913 / NCPB 528;
RX MEDLINE=22022145; PubMed=12024217;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Montelito-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Camarotte G., Camaravan F., Cardozo J., Chambergro F., Ciapina L.P.,
RA Ciarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locati E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spindola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities".
RL Nature 417:459-463(2002).
RN [2]
RP SEQUENCE OF 1-134 FROM N.A.
RC STRAIN=Xc1701;
RX MEDLINE=9641476; PubMed=8817497;
RA Hu N., Lee P., Chen C.;
RT "The type IV pre-pilin leader peptidase of Xanthomonas campestris pv.
RT campestris is functional without conserved cysteine residues.";
RL Mol. Microbiol. 18:769-777(1995).
CC -I- FUNCTION: CATALYZES THE PHOSPHORYLATION OF THE 3'-HYDROXYL GROUP

CC OF DEPHOSPHOCOENZYME A TO FORM COENZYME A (BY SIMILARITY).
CC -I- CATALYTIC ACTIVITY: ATP + dephospho-CoA -> ADP + CoA.
CC -I- PATHWAY: Coenzyme A (CoA) biosynthesis; fifth (last) step.
CC -I- SUBCELLULAR LOCATION: Cytoplasmic (by similarity).
CC -I- SIMILARITY: BELONGS TO THE COAE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE012425; AAM42373.1; -
DR EMBL; U12432; AAC43572.1; -
DR InterPro; IPR001977; DepP-CoAkinase.
DR Pfam; PF01121; Coae; 1.
DR ProDom; PD003329; DepP-CoAkinase; 1.
DR PROSITE; PS01294; COAE; 1.
KW Transferase; Kinase; ATP-binding; Coenzyme A biosynthesis.
FT NP_BIND 10 17
SQ SEQUENCE 207 AA; 22098 MW; 185F08185D1A1FD7 CRC64;

Query Match 100.0%; Score 21; DB 1; Length 207;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HARL 4
DB 141 HARL 144

RESULT 21
COAE_XYFPA STANDARD; PRT; 207 AA.
AC Q9P412;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Dephospho-CoA kinase (EC 2.7.1.24) (Dephosphocoenzyme A kinase).
GN COAE OR XF2536.
OS Xylella fastidiosa.
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
CC Xylella.
CX NCBI_TaxID=2371;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=9a5c;
RX MEDLINE=20365717; PubMed=10910347;
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA Alvarenga R., Alves L.M.C., Arya J.E., Bala G.S., Baptista C.S.,
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Britones M.R.S.,
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carter H.,
RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA Fraga J.S., Franca S.C., Franco M.C., Fromme M., Furlan L.R.,
RA Garner M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA Ho P.L., Hohnsels J.D., Junqueira M.L.S., Kemper E.L., Kitajima J.P.,
RA Krieger J.E., Kuramae E.B., Laigret F., Lambais M.R., Leite L.C.C.,
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Martino C.L.,
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Montelito-Vitorello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA Nhani A.Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,

```

RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsubako M.H.,
RA Vallada H., Van Sluys M.A., Verjovsky-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Meidanis J., Setubal J.C., Vettore A.L.,
RT "The genome sequence of the plant pathogen Xylella fastidiosa."
RL Nature 406:151-159(2000).
CC -1- FUNCTION: CATALYZES THE PHOSPHORYLATION OF THE 3'-HYDROXYL GROUP
CC OF DEPHOSPHOCOENZYME A TO FORM COENZYME A (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: ATP + dephospho-CoA -> ADP + CoA.
CC -1- PATHWAY: Coenzyme A (CoA) biosynthesis; fifth (last) step.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE COAE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE004060; AAF85333.1; ALT_INT.
DR InterPro: IPR001977; DepCoCkKinase.
DR Pfam: PF01121; Coae; 1.
DR ProDom: PD003329; DepCoCkKinase; 1.
DR TIGRFAMs: TIGR00152; UPF0038; 1.
DR PROSITE: PS01294; COAE; 1.
DR Transferase: Kinase; ATP-binding; Coenzyme A biosynthesis;
KW Complete proteome.
FT NP_BIND 10 17 ATP (POTENTIAL).
SQ SEQUENCE 207 AA; 22224 MW; FB6234199F314CD8 CRC64;

Query Match 100.0%; Score 21; DB 1; Length 207;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HARL 4
DB 141 HARL 144

RESULT 22
FUSD BURCE STANDARD; PRT; 208 AA.
AC P24129;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Fusaric acid resistance protein fUSD.
GN FUSD.
OS Burkholderia cepacia (Pseudomonas cepacia).
OC Bacteria; Proteobacteria; beta subdivision; Burkholderia group;
OC Burkholderia.
OX NCBI_TaxID=292;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OK1;
RX MEDLINE=9210974; PubMed=1370369;
RA Uysal R., Yagci T., Katayama S., Katsuragi K., Tachibana K.,
RA Toyoda H., Uchi S., Obata K., Shibano Y., Noda M.;
RT "Molecular cloning and characterization of the fusaric
RT acid-resistance gene from Pseudomonas cepacia."
RL Agric. Biol. Chem. 55:1913-1918(1991).
CC -1- FUNCTION: INVOLVED IN THE RESISTANCE (DETOXIFICATION) OF THE
CC FUNGAL TOXIN FUSARIC ACID.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

```

DR EMBL: S77489; AAC60391.1; -
DR EMBL: D12503; BAA02067.1; -
DR PIR: JS0512; JS0512.
KW Transmembrane.
FT TRANSMEM 7
SQ SEQUENCE 208 AA; 23039 MW; 50CDFE97489444E CRC64;

Query Match 100.0%; Score 21; DB 1; Length 208;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HARL 4
DB 46 HARL 49

RESULT 23
KTHY_AERPE STANDARD; PRT; 208 AA.
ID KTHY_AERPE
AC O9YA48;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable thymidylate kinase (EC 2.7.4.9) (dTMP kinase).
GN TMK OR APE2090.
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcaceae;
OC Desulfurococcaceae; Aeropyrum.
OX NCBI_TaxID=56636;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K1;
RX MEDLINE=99310339; PubMed=10382966;
RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Akai A., Kosugi H.,
RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA Yamazaki J., Kushida N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic
RT crenarchaeon, Aeropyrum pernix K1."
RL DNA Res. 6:83-101(1999).
CC -1- CATALYTIC ACTIVITY: ATP + thymidine 5'-phosphate -> ADP + thymidine
CC 5'-diphosphate.
CC -1- PATHWAY: SYNTHESIS OF dTMP FROM dTMP.
CC -1- SIMILARITY: BELONGS TO THE THYMIDYLATE KINASE FAMILY.
CC -1- CAUTION: COULD BE INACTIVE DUE TO A DEFECTIVE ATP-BINDING SITE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF000063; BAA8101.1; -
DR InterPro: IPR000062; Thymidylate_kin.
DR Pfam: PF02223; Thymidylate_kin; 1.
DR TIGRFAMs: TIGR00041; dTMP_kinase; 1.
DR PROSITE: PS01331; THYMIDYLATE_KINASE; 1.
DR Transferase: Kinase; Nucleotide biosynthesis; ATP-binding;
KW Complete proteome.
FT SITE 8 15 DEFECTIVE ATP-BINDING (POTENTIAL).
FT SITE SEQUENCE 208 AA; 22945 MW; ED155CE4EA3382F CRC64;

Query Match 100.0%; Score 21; DB 1; Length 208;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HARL 4
DB 22 HARL 25

```

RESULT 24
DKK4_HUMAN STANDARD; PRT; 224 AA.
AC Q9UBT3; O9Y4C3;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE DICKKOPF related protein-4 precursor (Dkk-4) (Dkk-4).
GN DKK4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP MEDLINE-20035735; PubMed-10570958;
RX MEDLINE-20035735; PubMed-10570958;
RA Krupnik V.E., Sharp J.D., Jiang C., Robison K., Chickering T.W.,
RA Amaravadi L., Brown D.E., Guyot D., Mays G., Leidy K., Chang B.,
RA Duong T., Goodenrich A.D., Gearing D.P., Sokol S.Y., McCarthy S.A.;
RT "Functional and structural diversity of the human Dkk-4 gene
family.";
RT Gene 238:301-313(1999).
RL [2]
RN SEQUENCE FROM N.A.
RP "Tate G., Mitsuya T.;
RT "Human Dkk-4 as well as DAN family members, Cerberus and Gremlin,
RT are preferentially expressed in the epithelial malignant cell lines.";
RL J. Biochem. Mol. Biol. Biophys. 3:239-242(1999).
RN [3]
RP SEQUENCE FROM N.A.
RA Tate G., Suzuki T., Mitsuya T.;
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: INHIBITOR OF WNT SIGNALING PATHWAY.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN CEREBELLUM, T-CELLS, ESOPHAGUS,
CC AND LUNG.
CC -1- PTM: APPEARS NOT TO BE GLYCOSYLATED.
CC -1- PTM: CAN ALSO BE PROTEOLYTICALLY PROCESSED BY A FURIN-LIKE
CC PROTEASE.
CC -1- SIMILARITY: BELONGS TO THE DICKKOPF FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF177397; AAF02677.1; -;
DR EMBL: AB018005; BAA33475.1; -;
DR EMBL: AB018003; BAA33475.1; JOINED.
DR EMBL: AB018004; BAA33475.1; JOINED.
DR EMBL: AB017788; BAA33438.1; -;
DR HSSP: P25687; 11MT.
DR GeneW: HGNC:2894; DKK4.
DR MIM: 605417; -;
KW Developmental protein; Signal.
FT SIGNAL 1 18
FT CHAIN 19 224 DICKKOPF RELATED PROTEIN-4.
FT CHAIN 224 224 DICKKOPF RELATED PROTEIN-4 SHORT FORM.
FT DOMAIN 41 90 DKK-TYPE CYS-1.
FT DOMAIN 145 218 DKK-TYPE CYS-2.
FT CONFILCT 93 93 M -> L (IN REF. 3).
SQ SEQUENCE 224 AA; 24875 MW; 45F8BEC476961357 CRC64;

Query Match 100.0%; Score 21; DB 1; Length 224;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4

DB 212 HARL 215
1111
RESULT 25
BIOD_XYLFA STANDARD; PRT; 226 AA.
AC Q9PAL9;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Dethiodiotin synthetase (EC 6.3.3.3) (Dethiodiotin synthetase) (DIB
DE synthetase) (DIBS).
GN BIOD OR XF2477.
OS Xylella fastidiosa.
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
OC Xylella.
OX NCBI_TaxID=2371;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-20365717; PubMed-10910347;
RA Simpson A.J.G., Rehnach F.C., Arruda P., Abreu F.A., Acencio M.,
RA Alvaranga R., Alves L.M.C., Araya J.B., Bala G.S., Baptista C.S.,
RA Barros M.H., Bonaccorsi E.D., Borlén S., Bove J.M., Bionnes M.R.S.,
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrer H.,
RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA Fraga J.S., Franco S.C., Franco M.C., Frohme M., Fulton L.R.,
RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA Ho P.L., Hobeisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA Krieger J.E., Kuramae E.E., Laigret F., Lampais M.R., Leite L.C.C.,
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
RA Marchais M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmeri D.A., Paris A.,
RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pezquerro J.B.,
RA Queglio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsubako M.H.,
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RL "The genome sequence of the plant pathogen Xylella fastidiosa.";
RL Nature 406:151-159(2000).
CC -1- CATALYTIC ACTIVITY: ATP + 7,8-diaminononanoate + CO(2) - ADP +
CC phosphate + dethiodiotin.
CC -1- COFACTOR: MAGNESIUM (BY SIMILARITY).
CC -1- PATHWAY: Bioconversion of pinelate into dethiodiotin.
CC -1- SIMILARITY: BELONGS TO THE DETHIODIOTIN SYNTHETASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE004056; AAF85275.1; -;
DR HSSP: P13000; 1DTS.
DR InterPro: IPR004472; BIOD.
DR InterPro: IPR002586; CblA_P.
DR Pfam: PF01656; CblA; 1.
DR TIGRFAMs: TIGR00347; BIOD; 1.
KW Biotin biosynthesis; ligase; Magnesium; ATP-binding;
KW Complete proteome.
FT NP_BIND 10 18 ATP (BY SIMILARITY).

```

SO SEQUENCE 226 AA; 23910 MW; 0CCBC4842A9BFBEE CRC64;
Query Match 100.0%; Score 21; DB 1; Length 226;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
    ||||
DB 156 HARL 159

RESULT 26
YP2C_STAAU
ID YP2C_STAAU STANDARD; PRT; 236 AA.
AC P14503;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 01-JUL-1993 (Rel. 26, Last annotation update)
DE Hypothetical (Rel. 7 kDa protein.
OS Staphylococcus aureus.
OG Plasmid PC223.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=RN154;
RA Dreher J., Matzura H.;
RX MEDLINE=88189801; Pubmed=3128771;
RT "Replication control of the Staphylococcus aureus chloramphenicol
  resistance plasmids PC223 and pUB112 in Bacillus subtilis.";
RL Nucleic Acids Res. 16:2045-2062(1988).
RC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X12831; CAA31313.1; .
DR PIR; S28101; S28101.
KW Hypothetical protein; plasmid.
SQ SEQUENCE 236 AA; 27676 MW; 206ED04E169EF34C CRC64;

Query Match 100.0%; Score 21; DB 1; Length 236;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
    ||||
DB 107 HARL 110

RESULT 27
HUTC_KLEAE
ID HUTC_KLEAE STANDARD; PRT; 241 AA.
AC P12380;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-FEB-1991 (Rel. 17, Last annotation update)
DE Histidine utilization repressor.
GN HUTC.
OS Klebsiella aerogenes.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Klebsiella.
OX NCBI_TaxID=28451;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RA MEDLINE=90368611; PubMed=2203754.;
RX Schwacha A., Bender R.A.;
RT "Nucleotide sequence of the gene encoding the repressor for the
RL histidine utilization genes of Klebsiella aerogenes.";
RN J. Bacteriol. 172:5477-5481(1990).
RN [2]
RN RP SEQUENCE OF 216-241 FROM N.A.
RX MEDLINE=86198018; PubMed=2834335;
RT Nieuwkoop A.J., Baldauf S.A., Hudspeth M.E.S., Bender R.A.;
RT "Bidirectional promoter in the hut(p) region of the histidine
RL utilization (hut) operons from Klebsiella aerogenes.";
RN J. Bacteriol. 170:2240-2246(1988).
CC -1- FUNCTION: THE HUTC PROTEIN IS A REPRESSOR WHICH BINDS TO THE
CC HTTP REGION IN THE HISTIDINE UTILIZATION (HUT) OPERON. IT BLOCKS
CC EXPRESSION OF ALL THE HUT GENES IN THE ABSENCE OF INDUCER.
CC -1- SIMILARITY: BELONGS TO THE GNR FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC in the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M34604; AAA25075.1; -;
DR EMBL; M19665; AAA25077.1; -;
DR PIR; B36730; B36730.
DR InterPro; IPR000524; HTH_Gntr.
DR Pfam; PF00392; gntr; 1.
DR PRINTS; PR00035; HTHGNTR.
DR SMART; SM00345; HTH_GNTR; 1.
DR PROSITE; PS00043; HTH_GNTR_FAMILY; 1.
KW Histidine metabolism; Repressor; Transcription regulation;
KW DNA-binding.
FT DNA_BIND 39 58 H-T-H MOTIF (BY SIMILARITY).
SO SEQUENCE 241 AA; 27215 MW; 61D1AB625BEAFB3B CnC64;

Query Match 100.0%; Score 21; DB 1; Length 241;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HARL 4
Db 223 HARL 226

RESULT 28
ID YT31_RALSO STANDARD; PRT; 248 AA.
AC Q8XVA0;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein RSC2931.
GN RSC2931 OR RS00159.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
OC Ralstonia.
OX NCBI_Taxid=305;
RN [1]
RN RP SEQUENCE FROM N.A.
RC STRAIN=GM1100;
RX MEDLINE=21681879; PubMed=11823852;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Ariat M., Billault A., Brottier P., Camu J.C., Cattolico L.,
RA Chardier M., Choise N., Claudel-Renard C., Cunac S., Demange N.,
RA Gaspin C., Lavie M., Moisan A., Robert C., Sautin W., Schlex T.,
RA Sigulier P., Thebault P., Whalen M., Winkler P., Levy M.,
RA Weissenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL Nature 415:497-502(2002).

```

```

CC -1- SIMILARITY: BELONGS TO THE UPF0135 (NIF3) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AL646072; CAD16638.1; -
DR InterPro: IPR002678; DUF34.
DR Pfam: PF01784; DUF34.1.
DR TIGRfams: TIGR00486; DUF34.1.
KM Hypothetical protein; Complete proteome.
SQ SEQUENCE 248 AA; 26651 MW; 101E0B25FC353F67 CRC64;

Query Match          100.0%; Score 21; DB 1; Length 248;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HARL 4
Db 80 HARL 83

RESULT 29
CBA4_LYCES STANDARD; PRT; 256 AA.
AC P27524;
DT 01-AUG-1992 (Rel. 23, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Chlorophyll A-B binding protein CP24 10A, chloroplast precursor (CAB-
DE 10A) (LHCP).
GN CAP10A.
OS Lycopersicon esculentum (tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-91355849; PubMed-2103430;
RA Schwartz E., Pichersky E.;
RT "Sequence of two tomato nuclear genes encoding chlorophyll
RT a/b-binding proteins of CP24, a PSI antenna component.";
RL Plant Mol. Biol. 15:157-160(1990).
CC -1- SIMILARITY: BELONGS TO THE EDIP/PSBS FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M32605; AAA34143.1; -
DR PIR: S11877; S11877.
DR InterPro: IPR001344; Chloro_Abbind.
DR Pfam: PF00504; Chloro_A-bind.1.
DR ProDom: PD000275; Chloro_Abbind.1.
KM Chlorophyll1; Photosynthesis; Photosystem II; Thylakoid;
KM Chloroplast; Transit peptide; Multigene family; Transmembrane.
FT TRANSIT 1 ?
FT CHAIN 256 ?
FT CHLOROPHYLL A-B BINDING PROTEIN CP24 10A.
SQ SEQUENCE 256 AA; 27253 MW; BD397BEF9E48300A CRC64;

Query Match          100.0%; Score 21; DB 1; Length 256;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 HARL 4
Db 229 HARL 232

RESULT 30
KC2C_YEAST STANDARD; PRT; 258 AA.
ID KC2C_YEAST
AC P38930;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Casein kinase II beta' chain (CK II).
GN CKB2 OR YOR039W OR OR26.32.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-94299539; PubMed-8027080;
RA Reed J.C., Bidwai A.P., Glover C.V.C.;
RT "Cloning and disruption of CKB2, the gene encoding the 32-kDa
RT regulatory beta'-subunit of Saccharomyces cerevisiae casein kinase
RT II.";
RL J. Biol. Chem. 269:18192-18200(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN-S288C / FY1679;
RA de Haan M., Maarse A.C., Grivell L.A.;
RL Submitted (MAY-1995) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Landt O., Hiesel R., Unsel M.;
RL Submitted (JUL-1996) to the EMBL/Genbank/DBJ databases.
RN [4]
RP SEQUENCE OF 172-258 FROM N.A.
RC STRAIN-AP3;
RA Haider M., Bito A., Wallner J., Breitenbach M.;
RL Submitted (JAN-1997) to the EMBL/Genbank/DBJ databases.
RN [5]
RP PARTIAL SEQUENCE.
RX MEDLINE-94182950; PubMed-8135547;
RA Bidwai A.P., Reed J.C., Glover C.V.C.;
RT "Casein kinase II of Saccharomyces cerevisiae contains two distinct
RT regulatory subunits, beta and beta'.";
RL Arch. Biochem. Biophys. 309:348-355(1994).
CC -1- FUNCTION: PLAYS A COMPLEX ROLE IN REGULATING THE BASAL CATALYTIC
CC ACTIVITY OF THE ALPHA SUBUNIT (BY SIMILARITY).
CC -1- SUBUNIT: TETRAMER COMPOSED OF AN ALPHA CHAIN, AN ALPHA', ONE
CC BETA CHAIN AND ONE BETA' CHAIN.
CC -1- PTM: PHOSPHORYLATED BY ALPHA CHAIN (BY SIMILARITY).
CC -1- PTM: THE N-TERMINUS IS BLOCKED.
CC -1- SIMILARITY: BELONGS TO THE CASEIN KINASE 2 BETA CHAIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U08849; AAA21656.1; -
DR EMBL: X87331; CA60758.1; -
DR EMBL: Z74947; CA99229.1; -
DR EMBL: X82893; CA58064.1; -
DR SGD: S0005565; CKB2.
DR InterPro: IPR000704; CAS_kinase_II.
DR Pfam: PF01214; CK_II_beta.1.
DR PRINTS: PR00472; CASKNINASEII.
DR PROSITE: PS01101; CK2_BETA.1.
KM Transferase; Serine/threonine-protein kinase; Phosphorylation.

```


SQ SEQUENCE 258 AA; 29842 MW; D6060CB3D70B19A4 CRC64;
 Query Match 100.0%; Score 21; DB 1; Length 258;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
 DB 96 HARL 99

RESULT 31
 C4_SPIOI
 ID C4_SPIOI STANDARD; PRT; 261 AA.
 AC P36494;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Chlorophyll A-B binding protein CP24, chloroplast precursor.
 OS Spinacia oleracea (Spinach)
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Caryophyllales; Magnoliophyta; eudicotyledons; core eudicots;
 OC Caryophyllidae; Caryophyllales; Chenopodiaceae; Spinacia.
 OX NCBI_TaxID=3562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=cv. Monatol; TISSUE=leaf;
 RL Submitted (SEP-1993) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE ELIP/PSBS FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC
 CC EMBL: Z25886; CAA81105.1; -
 DR PIR: S40210; S40210.
 DR InterPro: IPR001344; Chloro_Abbind.
 DR Pfam: PF00504; ChloroA-b-bind. 1.
 DR ProDom: PD000275; Chloro_Abbind. 1.
 KW Chlorophyll; Photosynthesis; Photosystem II; Thylakoid;
 KW Chloroplast; Transit peptide; Transmembrane.
 FT TRANSIT 1 51
 FT CHAIN 52 261 CHLOROPHYLL A-B BINDING PROTEIN CP24.
 SO SEQUENCE 261 AA; 27840 MW; A1599F7B626CA1A5 CRC64;

Query Match 100.0%; Score 21; DB 1; Length 261;
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
 DB 234 HARL 237

RESULT 32
 NIFM_KLEPN
 ID NIFM_KLEPN STANDARD; PRT; 266 AA.
 AC P08534;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-AUG-1988 (Rel. 08, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Protein nifM.
 OS NifM.
 GN Klebsiella pneumoniae.
 OS Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Klebsiella.
 OX NCBI_TaxID=573;
 RN [1]

RP SEQUENCE FROM N.A.
 RA MEDLINE=89094839; PubMed=3062178;
 RX Arnold W., Rump A., Klipp W., Pfeifer U.B., Puenhler A.;
 RT "Nucleotide sequence of a 24,206-base-pair DNA fragment carrying the
 RT entire nitrogen fixation gene cluster of Klebsiella pneumoniae";
 RL J. Mol. Biol. 203:715-738(1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=M5a1;
 RX MEDLINE=88082825; PubMed=2826155;
 RA Paul W., Merrick M.J.;
 RT "The nucleotide sequence of the nifH gene of Klebsiella pneumoniae
 RT and identification of a new nif gene: nifZ";
 RL Eur. J. Biochem. 170:259-265(1987).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA COLLET T.A., White T., Howard K., Orme-Johnson W.H.;
 RL Submitted (JUL-1989) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: REQUIRED FOR THE ACTIVATION AND STABILIZATION OF THE
 CC IRON-COMPONENT (NIFH) OF NITROGENASE. PROBABLE PRASE.
 CC -----
 CC -1- SIMILARITY: BELONGS TO THE PPIC/PARVULIN FAMILY OF ROTAMASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC
 CC EMBL: X05887; CAA29313.1; -
 DR EMBL: X13303; CAA31679.1; -
 DR EMBL: M24106; AAA25105.1; -
 DR PIR: S02510; S02510.
 DR InterPro: IPR000297; Rotamase.
 DR Pfam: PF00639; Rotamase; 1.
 DR PROSITE: PS01096; PPIC_PP1ASE_1; 1.
 DR PROSITE: PS01098; PPIC_PP1ASE_2; 1.
 KW Nitrogen fixation; Isomerase; Rotamase.
 FT DOMAIN 124 221
 SO SEQUENCE 266 AA; 30612 MW; CF51684ED161AAB CRC64;

Query Match 100.0%; Score 21; DB 1; Length 266;
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
 DB 88 HARL 91

RESULT 33
 TP32_TREPA
 ID TP32_TREPA STANDARD; PRT; 268 AA.
 AC O07950;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Membrane lipoprotein Tpn32 precursor (29 kDa protein).
 GN Tpn32 OR TP0821.
 OS Treponema pallidum.
 OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
 OX NCBI_TaxID=160;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Porcella S.F., Radolf J.D., Norgard M.V.;
 RT "Treponema pallidum 29k protein is homologous to a lipoprotein
 RT present in Pasteurella hemolytica and in Haemophilus influenzae
 RT type b.";
 RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.


```

RC STRAIN-Nichols;
RA Stramm L.V., Barnes N.Y.;
RT "Treponema pallidum lipoprotein homologue.";
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-Nichols;
RX MEDLINE=98332770; PubMed=9665876;
RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
RA Dodson R., Gwin M., Hickey E.K., Clayton R., Ketchum K.A.,
RA Dodegrem E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
RA Kralak H., Richardson D., Howell J.K., Chidambaram M., Utterback T.,
RA McDonald L., Attiach P., Bowman C., Cotton M.D., Fujii C., Garland S.,
RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
RA Venter J.C.;
RT "Complete genome sequence of Treponema pallidum, the syphilis
RT spirochete.";
RL Science 281:375-388(1998).
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
CC (Probable).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U97358; AAB63362.1; -
DR EMBL: U93844; AAB61267.1; -
DR EMBL: AE001253; AAC65789.1; -
DR TIGR: TP0821; -
DR InterPro: IPR004872; Lipoprotein_9.
DR Pfam: PF03180; Lipoprotein_9; 1.
DR PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
KM Membrane; Lipoprotein; Signal; Complete proteome.
PT SIGNAL 1 23 BY SIMILARITY.
FT CHAIN 24 268 MEMBRANE LIPOPROTEIN TP032.
FT LIPID 24 24 N-ACYL DIGLYCERIDE (POTENTIAL).
SQ SEQUENCE 268 AA; 29081 MW; AB21F8A07A88F691 CRC64;

Query Match 100.0%; Score 21; DB 1; Length 268;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
DB 41 HARL 44

RESULT 34
T2S1_STRFI STANDARD; PRT; 269 AA.
AC 052512;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Type II restriction enzyme SfiI (EC 3.1.21.4) (Endonuclease SfiI)
DE (R.SfiI).
GN SFIIR.
OS Streptomyces flambriatus.
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=68197;
RN [1]
RP SEQUENCE FROM N.A.
RA van Cott E.M., Moran L.S., Slatko B.E., Wilson G.G.;
RT "Characterization of the SfiI restriction and modification genes.";
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Recognizes the double-stranded sequence GGCCNNNNNGGCC
CC and cleaves before N-9.
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage of DNA to give

```

```

CC specific double-stranded fragments with terminal 5'-phosphates.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF039750; AAB95365.1; -
DR REBASE: 1655; SfiI.
KM Hydrolyase; Endonuclease; Nuclease; Restriction system.
SQ SEQUENCE 269 AA; 31044 MW; 3C48499BA5205EA CRC64;

Query Match 100.0%; Score 21; DB 1; Length 269;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
DB 188 HARL 191

RESULT 35
IC44_HUMAN STANDARD; PRT; 271 AA.
AC 014773; 014771; 014772; 016375;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Intercellular adhesion molecule-4 precursor (ICAM-4) (Landsteiner-
DE Wiener blood group glycoprotein) (LW blood group protein) (CD242
DE antigen).
GN ICAM4 OR LW.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94261574; PubMed=8202485;
RA Bailly P., Hermand P., Callebaut I., Sonneborn H.H., Khamilich S.,
RA Moron J.-P., Cartron J.-P.;
RT "The LW blood group glycoprotein is homologous to intercellular
RT adhesion molecules.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:5306-5310(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96219993; PubMed=8639917;
RA Hermand P., le Pennec P.Y., Rouger P., Cartron J.-P., Bailly P.;
RT "Characterization of the gene encoding the human LW blood group
RT protein in LW+ and LW- phenotypes.";
RL Blood 87:2962-2967(1996).
RN [3]
RP SEQUENCE OF 1-130 FROM N.A., AND VARIANT BLOOD GROUP LW(B).
RX MEDLINE=95359462; PubMed=7632968;
RA Hermand P., Gane P., Mattel M.-G., Sistonen P., Cartron J.-P.,
RA Bailly P.;
RT "Molecular basis and expression of the LWa/LWb blood group
RT polymorphism.";
RL Blood 86:1590-1594(1995).
CC -1- FUNCTION: ADHESION MOLECULE THAT BINDS TO LEUKOCYTE ADHESION
CC LFA-1 PROTEIN (INTEGRIN ALPHA-L/BETA-2).
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. THE SHORT FORM
CC LACKS A TRANSMEMBRANE DOMAIN AND IS EITHER SECRETED OR RAPIDLY
CC DEGRADED.
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS, A LONG FORM (SHOWN HERE) AND A
CC SHORT FORM. ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: ERYTHROCYTES.
CC -1- PTM: N- AND O-GLYCOSYLATED.
CC -1- POLYMORPHISM: LW IS RESPONSIBLE FOR THE LANDSTEINER-WIENER BLOOD
CC GROUP SYSTEM. THE MOLECULAR BASIS OF THE LW(A)-LW5/LW(B)-LW7 BLOOD

```

```

CC GROUP ANTIGENS IS A SINGLE VARIATION IN POSITION 100; GLN-100
CC CORRESPONDS TO LM(A) AND ARG-100 TO LM(B).
CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. ICAM
CC SUPERFAMILY.
CC -1- SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; L27671; AAA59538.1; -
DR EMBL; L27670; AAA59537.1; -
DR EMBL; X93093; CAA63646.1; -
DR EMBL; S78852; AAB35046.1; -
DR HSP; F05362; IIC1.
DR Genew; HGNC:5347; ICAM4.
DR MIM; 111250; -
KM Immunoglobulin domain: Cell adhesion; Glycoprotein; Transmembrane;
KW Repeat; Signal; Alternative splicing; Polymorphism;
FT BLOOD group antigen.
FT SIGNAL 1 22
FT CHAIN 23 271
FT DOMAIN 23 240
FT TRANSMEM 241 261
FT DOMAIN 262 271
FT DOMAIN 146 217
FT DOMAIN 146 217
FT DISULFID 153 210
FT DISULFID 68 78
FT CARBOHYD 78 68
FT CARBOHYD 190 190
FT CARBOHYD 223 223
FT VARSPLIC 233 271
FT VARIANT 100 100
FT CONFLICT 14 29
FT SEQUENCE 271 AA; 29265 MW; 2F6BC8BABD79E615 CRC64;
SQ
Query Match 100.0%; Score 21; DB 1; Length 271;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 HARL 4
DB 211 HARL 214

```

```

RT "Primary structure of the nuclear-encoded 29.9 kDa subunit of NADH:
RT ubiquinone reductase from Neurospora crassa mitochondria.";
RL Biochim. Biophys. Acta 1089:389-390(1991).
CC -1- FUNCTION: TRANSFER OF ELECTRONS FROM NADH TO THE RESPIRATORY
CC CHAIN. THE IMMEDIATE ELECTRON ACCEPTOR FOR THE ENZYME IS BELIEVED
CC TO BE UBIQUINONE.
CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC -1- CATALYTIC ACTIVITY: NADH + acceptor = NAD(+) + reduced acceptor.
CC -1- SUBUNIT: COMPLEX I IS COMPOSED OF ABOUT 30 DIFFERENT SUBUNITS.
CC -1- SUBCELLULAR LOCATION: Mitochondrial inner membrane; matrix side.
CC -1- SIMILARITY: BELONGS TO THE COMPLEX I 29.9 kDa SUBUNIT FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; X56237; CAA39694.1; -
DR PIR; S17191; S17191.
KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion; Transil peptide.
FT TRANSIT 1 8
FT CHAIN 9 273
FT SEQUENCE 273 AA; 30862 MW; 4B1175993B47EEF1 CRC64;
SQ
Query Match 100.0%; Score 21; DB 1; Length 273;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 HARL 4
DB 113 HARL 116

```

```

CC -----
DR EMBL: U57906; AAB49191.1; -
DR HSSP: P03017; 2REB.
DR InterPro: IPR001553; RECA.
DR Pfam: PF00154; RECA.1.
DR ProDom: PD000229; RECA.1.
DR PROSITE: PS00321; RECA_1; 1.
DR PROSITE: PS50162; RECA_2; 1.
DR PROSITE: PS50163; RECA_3; 1.
DR DNA damage; DNA recombination; SOS response; ATP-binding; DNA-binding.
FT NON_TER 1
FT 43 50
FT NP_BIND 274
FT NON_TER 274
SQ SEQUENCE 274 AA; 29602 MW; 4EF5C1097339D94 CRC64;

Query Match
Best Local Similarity 100.0%; Score 21; DB 1; Length 274;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
Db 144 HARL 147

RESULT 38
YG73_SYNY3
ID YG73_SYNY3 STANDARD; PRT; 274 AA.
AC P14261;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical tRNA/tRNA methyltransferase Slr1673 (EC 2.1.1.-).
GN Slr1673.
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97061201; PubMed=8905231;
RA Kaneo T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hirosewa M., Sugita A., Sasamoto S., Kimura T.,
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA Shimizu S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA Tabata S.;
RT *Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.;
RL DNA Res. 3:109-136(1996).
CC -1- SIMILARITY: BELONGS TO THE RNA METHYLTRANSFERASE TRNA FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D90913; BAA18355.1; -
DR InterPro: IPR001537; Spou_methylase.
DR Pfam: PF00568; Spou_methylase; 1.
KW Hypothetical protein; Transferase; Methyltransferase;
KW Complete proteome.
SQ SEQUENCE 274 AA; 29701 MW; A4176C1061CAE88 CRC64;

Query Match
Best Local Similarity 100.0%; Score 21; DB 1; Length 274;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
Db 118 HARL 121

```

```

RESULT 39
CTA2_ACT1M
ID CTA2_ACT1M STANDARD; PRT; 275 AA.
AC O33950;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Catechol 1,2-dioxygenase 2 (EC 1.13.11.1) (CD12).
GN CTA2.
OS Acinetobacter lwoffii.
OC Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;
OC Acinetobacter.
OX NCBI_TaxID=28090;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-30; 27-65 AND 229-247.
RC STRAIN-K24;
RX MEDLINE=97405925; PubMed=9260969;
RA Kim S.T., Leem S.-H., Choi J.-S., Chung Y.H., Kim S., Park Y.-M.,
RA Park Y.K., Lee Y.N., Ha K.-S.;
RT "Cloning and characterization of two catA genes in Acinetobacter
RT lwoffii K24.";
RL J. Bacteriol. 179:5226-5231(1997).
CC -1- FUNCTION: CAN CLEAVE 4-METHYL-, 4-CHLORO-, AND 3-METHOXYCATECHOL
CC AT LOWER RATES THAN CATECHOL, BUT HAS NO ACTIVITY WITH 4-
CC NITROCATATECHOL OR PROTOCATECHOLIC ACID.
CC -1- CATALYTIC ACTIVITY: Catechol + O(2) -> cis,cis-muconate.
CC -1- COFACTOR: FERRIC ION.
CC -1- PATHWAY: FIRST STEP IN THE DEGRADATION OF CATECHOL TO SUCCINATE
CC AND ACETYL-COA IN THE BETA-KETOADIPATE PATHWAY.
CC -1- SUBUNIT: HOMODIMER (POTENTIAL).
CC -1- INDUCTION: BY ANILINE.
CC -1- SIMILARITY: BELONGS TO THE INTRADIOL RING-CLEAVAGE DIOXYGENASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U77659; AAC31767.1; -
DR HSSP: P00437; 3PCC.
DR InterPro: IPR000627; Dioxygenase.
DR Pfam: PF00775; Dioxygenase; 1.
DR PROSITE: PS00083; INTRADIOL_DIOXYGENAS; 1.
KW Aromatic hydrocarbons catabolism; Oxidoreductase; Dioxygenase; Iron.
FT METAL 158 158
FT METAL 192 192
FT METAL 216 216
FT METAL 218 218
SQ SEQUENCE 275 AA; 30399 MW; 3F8B0C127F45A807 CRC64;

Query Match
Best Local Similarity 100.0%; Score 21; DB 1; Length 275;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
Db 115 HARL 118

RESULT 40
DAF_PSEAE
ID DAF_PSEAE STANDARD; PRT; 276 AA.
AC O51564;
DT 01-NOV-1997 (Rel. 35, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Diaminopimelate epimerase (EC 5.1.1.7) (DAP epimerase).
GN DAPF OR PA5278.

```

```

OS Pseudomonas aeruginosa.
OC Bacteria: Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas
OX NCBI_TaxId=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 15692 / PAOI;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Huftagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Lardys K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT Complete genome sequence of Pseudomonas aeruginosa PAOI, an
RT opportunistic pathogen.;
RL Nature 406:959-964(2000).
RN [2]
RP SEQUENCE OF 169-276 FROM N.A.
RC STRAIN=7NSK2;
RX MEDLINE=95231283; PubMed=7715441;
RA Hofte M., Dong O., Kourambas S., Krishnapillai V., Mergely M.;
RT The sss gene product, which affects pyoverdinin production in
RT Pseudomonas aeruginosa 7NSK2, is a site-specific recombinase.;
RL Mol. Microbiol. 14:1011-1020(1994).
CC -1- CATALYTIC ACTIVITY: L1-2,6-diaminohexanoate - meso-
CC diaminoheptanedioate.
CC -1- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF LYSINE FROM ASPARTATE
CC SEMIALDEHYDE.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE DIAMINOPELATE EPIMERASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE004940; AAG08663.1; -
DR EMBL; X78478; CAA55224.1; -
DR HSSP; P44859; IBM2.
DR InterPro: IPR001653; DAP-epimerase.
DR Pfam: PF01678; DAP-epimerase; 2.
DR TIGRfams: TIGR00652; DapF; 1.
DR PROSITE: PS01326; DAP-EPIMERASE; 1.
KM Isomerase; lysine biosynthesis; Complete proteome.
FT ACT_SITE 75 75 BY SIMILARITY.
FT ACT_SITE 219 219 BY SIMILARITY.
SQ SEQUENCE 276 AA; 30320 MW; 42833EE0611621F CRC64;

Query Match 100.0%; Score 21; DB 1; Length 276;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HARL 4
DB 203 HARL 206

```

```

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93219141; PubMed=8464751;
RA Winkler M., Rueger W.;
RT Cloning and sequencing of the genes of
RT beta-glucosyl-HMC-alpha-glucosyl-transferases of bacteriophages T2
RT and T6.;
RL Nucleic Acids Res. 21:1500-1500(1993).
CC -1- FUNCTION: TRANSFERS A GENTIOBIOSYL-GROUP ON AN
CC HYDROXYMETHYLCTOSINE RESIDUE IN DNA.
CC -1- PATHWAY: INVOLVED IN THE DNA MODIFICATION PROCESS TO PROTECTS
CC RESTRICTION ENDONUCLEASE SYSTEM.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X68724; CAA48664.1; -
DR PIR; S35623; S35623.
KM Transferase; Glycosyltransferase.
SQ SEQUENCE 280 AA; 32396 MW; 75789FEEF4C3C88E CRC64;

Query Match 100.0%; Score 21; DB 1; Length 280;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HARL 4
DB 128 HARL 131

RESULT 42
GSTRC_BPT6 STANDARD; PRT; 280 AA.
ID GSTRC_BPT6
AC 006718;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE Beta-glucosyl-HMC-alpha-glucosyl-transferase (EC 2.4.1.-).
OS Bacteriophage T6.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;
OC T4-like viruses.
OX NCBI_TaxId=10666;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93219141; PubMed=8464751;
RA Winkler M., Rueger W.;
RT Cloning and sequencing of the genes of
RT beta-glucosyl-HMC-alpha-glucosyl-transferases of bacteriophages T2
RT and T6.;
RL Nucleic Acids Res. 21:1500-1500(1993).
CC -1- FUNCTION: TRANSFERS A GENTIOBIOSYL-GROUP ON AN
CC HYDROXYMETHYLCTOSINE RESIDUE IN DNA.
CC -1- PATHWAY: INVOLVED IN THE DNA MODIFICATION PROCESS TO PROTECTS
CC THE PHAGE GENOME AGAINST ITS OWN NUCLEASES AND THE HOST
CC RESTRICTION ENDONUCLEASE SYSTEM.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X68725; CAA48667.1; -
DR PIR; S35626; S35626.
KM Transferase; Glycosyltransferase.

```

SQ SEQUENCE 280 AA; 32299 MW; 44AA142262658BDE CRC64;
 Query Match 100.0%; Score 21; DB 1; Length 280;
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 HARL 4
 Db 128 HARL 131
 RESULT 43
 MDCB_PSEPU STANDARD; PRT; 280 AA.
 AC 092452;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 15-OCT-2001 (Rel. 40, Last sequence update)
 DE 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Probable 2-(5'-triphosphoribosyl)-3'-dephosphocoenzyme-A synthase
 (EC 4.2.-.-) (2-(5'-triphosphoribosyl)-3'-dephospho-CoA synthase).
 GN MDCB.
 OS Pseudomonas putida.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 OX NCBI_TaxID=303;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=IAM 1177;
 RA MEDLINE=99271182; PubMed=10339824;
 RA Chohan S., Kurusu Y., Nishihara H., Takamura Y.;
 RT "Cloning and characterization of mdc genes encoding malonate
 decarboxylase from Pseudomonas putida."
 RL FEMS Microbiol. Lett. 174:311-319(1999).
 CC -1- FUNCTION: INVOLVED IN THE FORMATION OF 2-(5'-PHOSPHORIBOSYL)-3'-
 DEPHOSPHOCOENZYME-A, THE PROSTHETIC GROUP OF THE ACYL-CARRIER
 PROTEIN OF THE MALONATE DECARBOXYLASE.
 CC -1- CATALYTIC ACTIVITY: Dephospho-CoA + ATP -> 2-(5'-
 triphosphoribosyl)-3'-dephospho-CoA + adenine.
 CC -1- SIMILARITY: BELONGS TO THE CITG/MDCB FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AB017138; BAA36205.1; ALT_INT.
 DR InterPro: IPR002736; CITG.
 DR Pfam: PF01874; CITG; 1.
 KW Lyase.
 SQ SEQUENCE 280 AA; 29150 MW; F43340A7DEB975AD CRC64;
 Query Match 100.0%; Score 21; DB 1; Length 280;
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 HARL 4
 Db 195 HARL 198
 RESULT 44
 GPR40_HUMAN STANDARD; PRT; 300 AA.
 AC 014842;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Putative G protein-coupled receptor GPR40.
 GN GPR40.
 OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98008875; PubMed=9344866;
 RA Sawdargo M., George S.R., Nguyen T., Xu S., Kolakowski L.F. Jr.,
 RA O'Dowd B.F.;
 RT "A cluster of four novel human G protein-coupled receptor genes
 occurring in close proximity to CD22 gene on chromosome 19q13.1."
 RL Biochem. Biophys. Res. Commun. 239:543-547(1997).
 CC -1- FUNCTION: ORPHAN RECEPTOR.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF024687; AAB86710.1;
 DR GeneW; HGNC:4498; GPR40.
 DR MIM; 603820;
 DR InterPro: IPR00276; GPCR_Rhodpsn.
 DR Pfam: PF00001; 7tm.1; 1.
 DR PROSITE: PS00237; G-PROTEIN RECP_FL_1; FALSE_NEG.
 DR PROSITE: PS50262; G-PROTEIN RECP_FL_2; 1.
 KM G-protein coupled receptor; Transmembrane.
 FT DOMAIN 1 10
 FT TRANSMM 1 1
 FT DOMAIN 31 31
 FT TRANSMM 32 42
 FT DOMAIN 43 63
 FT TRANSMM 64 71
 FT DOMAIN 82 102
 FT TRANSMM 103 123
 FT DOMAIN 124 144
 FT TRANSMM 145 178
 FT DOMAIN 179 197
 FT TRANSMM 198 221
 FT TRANSMM 222 242
 FT DOMAIN 243 258
 FT TRANSMM 259 279
 FT DOMAIN 280 300
 FT CARBOHYD 155 155
 SQ SEQUENCE 300 AA; 31457 MW; 77EF27DADC93380B CRC64;
 Query Match 100.0%; Score 21; DB 1; Length 300;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 HARL 4
 Db 33 HARL 36
 RESULT 45
 MMSR_PSEAB STANDARD; PRT; 307 AA.
 AC P28809;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE MMSR operon regulatory protein.
 GN MMSR OR PA3571.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 OX NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.

```

RC STRAIN-ATCC 15692 / PA01;
RA MEDLINE-92317087; PubMed-1339433;
RA Steele M.I., Lorenz D., Hatter K., Park A., Sokatch J.R.;
RT "Characterization of the mmsAB operon of Pseudomonas aeruginosa PAO
RT encoding methylmalonate-semialdehyde dehydrogenase and 3-
RT hydroxyisobutyrate dehydrogenase";
RL J. Biol. Chem. 267:13585-13592(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 15692 / PA01;
RX MEDLINE-20437337; PubMed-10984043;
RA Steyer C.K., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Hickey M.J., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Garber R.L., Coltry L., Tolentino E., Westbrook-Wadman S., Yan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Laird K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.R.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen";
RL Nature 406:959-964(2000).
CC -!- FUNCTION: REGULATORY PROTEIN FOR THE MMSAB OPERON. ACTIVATES THE
CC TRANSCRIPTION OF THE MMSAB GENES.
CC -!- SIMILARITY: BELONGS TO THE ARAC/XYLIS FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M84911; AAA25890.1; -
DR EMBL; AE004778; AAG0659.1; -
DR PIR; S27601; S27601.
DR PIR; A42902; A42902.
DR InterPro; IPR000005; HTHARAC.
DR Pfam; PF00165; HTH_ARAC; 2.
DR PRINTS; PR00032; HTHARAC.
DR SMART; SM00342; HTH_ARAC; 1.
DR PROSITE; PS00041; HTH_ARAC_FAMILY_1; 1.
DR PROSITE; PS01124; HTH_ARAC_FAMILY_2; 1.
RT Transcription regulation; DNA-binding; Activator; Complete proteome.
FT DNA_BIND 217 236
FT CONFLICT 129 129 S -> T (IN REF. 1).
SQ SEQUENCE 307 AA; 35379 MW; 9BF7AE348FAFD45 CRC64;

```

```

Query Match 100.0%; Score 21; DB 1; Length 307;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 HARL 4
DB 212 HARL 215

```

```

RESULT 46
MAL3_SCHPO STANDARD; PRT; 308 AA.
ID MAL3_SCHPO
AC 010113;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Microtubule integrity protein mal3.
GN MAL3 OR SPAC1866.15.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomyces.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RX MEDLINE-98012204; PubMed-9348288;
RA Behnauer J.D., Hagan I.M., Hegemann J.H., Fleig U.;
RT "Mal3, the fission yeast homologue of the human APC-interacting
RT protein EB-1 is required for microtubule integrity and the
RT maintenance of cell form.";
RL J. Cell Biol. 139:717-728(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-972;
RX MEDLINE-21848401; PubMed-11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Brooks J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Sgouras K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holtroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule L., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volkert G., Aert R., Robben J., Grynolprez B.,
RA Wellens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lebrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wandut R., Furnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galbert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gallardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerritielli L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakowski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
CC -!- FUNCTION: MAY PLAY A ROLE IN REGULATING THE INTEGRITY OF
CC MICROTUBULES POSSIBLY BY INFLUENCING THEIR STABILITY.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC; ASSOCIATED WITH MICROTUBULES.
CC -!- SIMILARITY: TO YEAST BIM1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Y09518; CAAT0707.1; -
DR EMBL; Z68198; CAA92392.1; -
DR InterPro; IPR004953; EBI.
DR Pfam; PF03271; EBI; 1.
KW Microtubules.
SQ SEQUENCE 308 AA; 35094 MW; 9527D31CA024256 CRC64;

```

```

Query Match 100.0%; Score 21; DB 1; Length 308;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 HARL 4
DB 292 HARL 295

```

```

RESULT 47
MIAA_STRCO STANDARD; PRT; 312 AA.
ID MIAA_STRCO
AC 069967;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE tRNA delta(2)-isopentenylpyrophosphate transferase (EC 2.5.1.8) (1PP

```

DE transferase) (Isopentenyl-diphosphate:trna isopentenyltransferase)
 DE (IPPT)
 GN MRAA OR SC05791 OR SC4H2.12.
 OS Streptomyces coelicolor.
 OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
 OC Actinomycetales; Streptomyicinae; Streptomycetaceae; Streptomyces.
 RN NCBI_TaxID=1902;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2) / M145;
 RX MEDLINE=21996410; PubMed=12000953;
 RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleser H.,
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
 RA Cronin A., Fraser A., Gobie A., Hidalgo J., Hornsby T., Howarth S.,
 RA Huang C.-H., Kleser T., Larke L., Murphy L., Oliver K., O'Neill S.,
 RA Rabbittowitsch E., Rajadream M.A., Rutherford K., Rutter S.,
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
 RA Warren T., Wietzorek A., Woodward J., Barrell B.G., Parkhill J.,
 RA Hopwood D.A.;
 RT "Complete genome sequence of the model actinomycete Streptomyces
 RT coelicolor A3(2).";
 RL Nature 417:141-147(2002).
 CC -1- FUNCTION: CATALYZES THE FIRST STEP IN THE BIOSYNTHESIS OF
 CC 2-METHYLMETHIO-N6-(DELTA(2)-ISOPENTENYL)-ADENOSINE (MS[211(6)A])
 CC ADJACENT TO THE ANTICODON OF SEVERAL TRNA SPECIES (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: isopentenyl diphosphate + trna -> diphosphate +
 CC trna containing 6-isopentenyladenosine.
 CC -1- SIMILARITY: BELONGS TO THE IPP TRANSFERASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AL022268; CAA18328.1; -
 DR InterPro: IPR002627; IPPT.
 DR Pfam: PF01715; IPPT. 1.
 DR ProDom: PD004674; IPPT. 1.
 DR TIGRfam: TIGR00174; mlaa. 1.
 DR Trnaseq: Nucleotidyltransferase; trna processing; ATP-binding;
 KW Complete proteome.
 FT NP_BIND 15 22 ATP (POTENTIAL).
 SO SEQUENCE 312 AA; 33967 MW; 347DB32F4BB60C1B CRC64;
 OY Query Match 100.0%; Score 21; DB 1; Length 312;
 DB Best Local Similarity 100.0%; Pred. No. 2.1e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 HARL 4
 DB 143 HARL 146
 RESULT 48
 US10_HSV11 STANDARD; PRT; 312 AA.
 AC P06486;
 DT 01-JUN-1988 (Rel. 06, Created)
 DT 01-JUN-1988 (Rel. 06, Last sequence update)
 DT 01-APR-1993 (Rel. 25, Last annotation update)
 DE Virion protein US10.
 GN US10.
 OS Herpes simplex virus (type 1 / strain 17).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Simplexvirus.
 OX NCBI_TaxID=10299;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=85160822; PubMed=2984429;

RA McGeoch D.J., Dolan A., Donald S., Rixon F.J.;
 RT "Sequence determination and genetic content of the short unique
 RT region in the genome of herpes simplex virus type 1.";
 RL J. Mol. Biol. 181:1-13(1985).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=84169548; PubMed=6324121;
 RA Rixon F.J., McGeoch D.J.;
 RT "A 3' co-terminal family of mRNAs from the herpes simplex virus type
 RT 1 short region: two overlapping reading frames encode unrelated
 RT polypeptide one of which has highly reiterated amino acid sequence.";
 RL Nucleic Acids Res. 12:2473-2487(1984).
 CC -1- SIMILARITY: BELONGS TO A FAMILY THAT GROUP TOGETHER HSV-1 US10,
 CC EHV-1 66, EHV-4 ORF3, AND VZV 64/69.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X14112; CAA32275.1; -
 DR EMBL: X02138; CAA26064.2; -
 DR EMBL: X00428; CAA25126.1; -
 DR EMBL: L00036; AAA96678.1; -
 DR PIR: A05242; Q0BE07.
 DR InterPro: IPR000087; Collagen.
 DR InterPro: IPR000714; EHV_Unk.
 DR Pfam: PF02053; Gene66; 1.
 DR Zn_Finger.
 FT ZN_FING 271 293 POTENTIAL.
 SO SEQUENCE 312 AA; 34055 MW; 12D01B0E7C920EA3 CRC64;
 OY Query Match 100.0%; Score 21; DB 1; Length 312;
 DB Best Local Similarity 100.0%; Pred. No. 2.1e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 HARL 4
 DB 222 HARL 225
 RESULT 49
 YCHK_ECOLI STANDARD; PRT; 314 AA.
 ID YCHK_ECOLI
 AC P37053;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein ychk.
 GN YCHK OR B1234.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=94110230; PubMed=8282700;
 RA Boesl M., Kersten H.;
 RT "Organization and functions of genes in the upstream region of tyrt
 RT of Escherichia coli: phenotypes of mutants with partial deletion of a
 RT new gene (tgs).";
 RL J. Bacteriol. 176:221-231(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,


```

RA  Man B., Shao Y.;
RT  "The complete genome sequence of Escherichia coli K-12.";
RL  Science 277:1453-1474(1997).
RN  [3]
RP  SEQUENCE FROM N.A.
RC  STRAIN-K12;
RX  MEDLINE=97251357; PubMed=9097039;
RA  Alpha H., Baba T., Fujita K., Hayashi K., Inada T., Isono K.,
RA  Itoh T., Kasai H., Kashimoto K., Kimura S., Kitakawa M.,
RA  Kitagawa M., Makino K., Miki T., Mizobuchi K., Mori H., Mori T.,
RA  Motomura K., Nakade S., Nakamura Y., Nishimoto H., Nishio Y.,
RA  Oshima T., Saito J., Saito G., Seki Y., Sivasubraman S.,
RA  Tagami H., Takeda J., Takemoto K., Takeuchi Y., Wada C.,
RA  Yamamoto Y., Horiuchi T.;
RT  "A 570-kb DNA sequence of the Escherichia coli K-12 genome
RT  corresponding to the 28.0-40.1 min region on the linkage map.";
RL  DNA Res. 3:363-377(1996).
RN  [4]
RP  SEQUENCE FROM N.A.
RC  STRAIN-K12;
RX  MEDLINE=97061202; PubMed=8905232;
RA  Oshima T., Alpha H., Baba T., Fujita K., Hayashi K., Honjo A.,
RA  Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
RA  Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
RA  Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,
RA  Samped G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA  Yano M., Horiuchi T.;
RT  "A 718-kb DNA sequence of the Escherichia coli K-12 genome
RT  corresponding to the 12.7-28.0 min region on the linkage map.";
RL  DNA Res. 3:137-155(1996).
CC  -1- SIMILARITY: BELONGS TO THE UPF0028 (SMS) FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; M64675; -; NOT_ANNOTATED_CDS.
DR  EMBL; AE000221; AAC74316.1; -
DR  EMBL; D90852; BAA16028.1; -
DR  EMBL; D90758; BAA36102.1; -
DR  EMBL; D90759; BAA36114.1; -
DR  PIR; B36871; B36871.
DR  EcoGene; EG12120; ychK.
DR  InterPro: IPR002641; Patatin.
DR  InterPro: IPR001423; UPF0028.
DR  Pfam; PF01734; Patatin; 1.
DR  PROSITE; PS01237; UPF0028; 1.
KM  Hypothetical protein; Complete proteome.
SQ  SEQUENCE 314 AA; 34355 MW; 5ABBD877C72FE40 CRC64;

OY  1 HARL 4
    ||||
DB  223 HARL 226

Query Match      100.0%; Score 21; DB 1; Length 314;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OC  Bacteria; Thermotogae; Thermotogae (class); Thermotogales;
OX  NCBI_TaxID=2336;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN-MSB8 / DSM 3109;
RX  MEDLINE=99287316; PubMed=10360571;
RA  Nelson K.E., Clayton R.A., Gill S.R., Gwin M.L., Dodson R.J.,
RA  Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA  McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA  Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA  Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA  Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT  "Evidence for lateral gene transfer between Archaea and Bacteria from
RT  genome sequence of Thermotoga maritima.";
RL  Nature 399:323-329(1999).
CC  -1- SIMILARITY: BELONGS TO THE UPF0052 FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; AE001810; AAD36776.1; -
DR  TIGR; TM1709; -
DR  InterPro: IPR002882; UPF0052.
DR  Pfam; PF01933; UPF0052; 1.
KM  Hypothetical protein; Complete proteome.
SQ  SEQUENCE 314 AA; 34632 MW; 886FBE17F3E06F44 CRC64;

OY  1 HARL 4
    ||||
DB  126 HARL 129

Query Match      100.0%; Score 21; DB 1; Length 314;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Search completed: November 13, 2002, 13:41:17
Job time : 11 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2002, 13:39:48 ; Search time 29 Seconds

(without alignments)
28.420 Million cell updates/sec

Title: US-09-697-590-2_COPY_292_295

Perfect score: 21

Sequence: 1 HARL 4

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 segs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTRMBL_21.*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mmc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriaph:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	21	100.0	31	2	09F115
2	21	100.0	37	4	09BQ22
3	21	100.0	39	12	09IW57
4	21	100.0	42	10	09S083
5	21	100.0	44	16	08U546
6	21	100.0	45	16	08Y1N1
7	21	100.0	52	5	097152
8	21	100.0	57	16	08XUV9
9	21	100.0	59	17	027782
10	21	100.0	65	2	007026
11	21	100.0	67	16	08UJ23
12	21	100.0	68	12	08VAB5
13	21	100.0	69	4	09BU94
14	21	100.0	69	12	09YP21
15	21	100.0	70	10	0943U1
16	21	100.0	74	12	09IK63

17	21	100.0	74	16	033186	033186 mycobacteri
18	21	100.0	75	10	065817	065817 hotdenum vul
19	21	100.0	79	2	09XB07	09XB07 myxococcus
20	21	100.0	80	16	09PDG2	09PDG2 xylella fas
21	21	100.0	80	16	08XYV7	08XYV7 ralsstonia s
22	21	100.0	81	10	09M675	09M675 cucumis mel
23	21	100.0	83	12	09WGD5	09WGD5 human papill
24	21	100.0	83	16	09K5X7	09K5X7 bacillus ha
25	21	100.0	85	12	066074	066074 canine hept
26	21	100.0	85	16	0807C1	0807C1 agrobacteri
27	21	100.0	87	2	09FD17	09FD17 streptomyce
28	21	100.0	87	16	09FBZ3	09FBZ3 streptomyce
29	21	100.0	88	12	09YKD5	09YKD5 tomato yell
30	21	100.0	92	2	0935B6	0935B6 frankia sp.
31	21	100.0	93	5	018842	018842 caenorhabdi
32	21	100.0	93	16	0911I3	0911I3 pseudomonas
33	21	100.0	95	16	08XK19	08XK19 salmonella
34	21	100.0	96	9	038025	038025 bacterioph
35	21	100.0	97	10	09M9W0	09M9W0 arabidopsis
36	21	100.0	99	2	087620	087620 pseudomonas
37	21	100.0	100	16	0926Q7	0926Q7 rickettsia
38	21	100.0	101	2	0930I7	0930I7 bacteroides
39	21	100.0	101	16	09JRG6	09JRG6 neisseria m
40	21	100.0	102	12	037880	037880 tomato yell
41	21	100.0	103	12	09W838	09W838 kaposi's sa
42	21	100.0	103	16	08UG19	08UG19 agrobacteri
43	21	100.0	103	16	08U5L6	08U5L6 agrobacteri
44	21	100.0	103	16	08Y0F6	08Y0F6 ralsstonia s
45	21	100.0	104	11	09D6F0	09D6F0 mus musculus

ALIGNMENTS

RESULT 1	ID	09F115	PRELIMINARY;	PRT;	31 AA.
AC	09F115;				
DT	01-MAR-2001 (TREMBLrel. 16, Created)				
DT	01-MAR-2001 (TREMBLrel. 16, Last sequence update)				
DT	01-MAR-2001 (TREMBLrel. 16, Last annotation update)				
DE	Hypothetical 3.5 kDa protein.				
GN	EP0010.				
OS	Enterococcus faecalis (Streptococcus faecalis).				
OC	Plasmid PAM373.				
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;				
OC	Enterococcaceae; Enterococcus.				
OX	NCBI_TaxID=1351;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=20453452; PubMed=10998166;				
RA	De Boeyer E.H., Clewell D.B., Fraser C.M.;				
RT	"Enterococcus faecalis conjugative plasmid PAM373: complete nucleotide				
RT	sequence and genetic analyses of sex pheromone response.";				
RL	EMBL; AE002565; AAG40421.1; -				
DR	Hypothetical protein; Plasmid.				
KW	SEQUENCE 31 AA; 3509 MW; 4E19CB94B3DB9421 CRC64;				
SQ					
Query Match	Best Local Similarity	100.0%;	Score 21;	DB 2;	Length 31;
Matches	4; Conservative	0;	Pred. No. 1.9e+02;	Mismatches 0;	Indels 0; Gaps 0;
QY	1 HARL 4				
DB	12 HARL 15				
RESULT 2	ID	09BQ22	PRELIMINARY;	PRT;	37 AA.
AC	09BQ22;				
DT	01-JUN-2001 (TREMBLrel. 17, Created)				

```

DT 01-JUN-2001 (TREMBlrel. 17, last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, last annotation update)
DE Alpha-2-macroglobulin (Fragment).
GN A2M.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Telch N., Mossner J., Keim V.;
RT "Coding sequence of alpha-2-macroglobulin exon 18 and flanking
RT introns.";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Telch N., Mossner J., Keim V.;
RT "Homo sapiens alpha-2-macroglobulin (A2M), exon 18, A2M-2 allele.";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF349032; AAK38109.1; -
DR EMBL; AF349033; AAK38110.1; -
DR InterPro; IPR001599; MacroglublnA2.
DR Pfam; PF00207; A2M; 1.
FT NON_TER 37 37
FT NON_TER 1 1
SQ SEQUENCE 37 AA; 4390 MW; AFEIC0124ACBF415 CRC64;

Query Match
Best Local Similarity 100.0%; Score 21; DB 4; Length 37;
Pred. No. 2.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
DB 8 HARL 11

RESULT 3
OY1W57 PRELIMINARY; PRT; 39 AA.
ID O91W57
AC O91W57;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, last annotation update)
DE Glycoprotein (Fragment).
GN GP57.
OS Borna disease virus (BDV).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales; Bornaviridae;
OC Borna virus.
OC NCBI_TaxID=12455;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-RATBDV;
RX MEDLINE=20086018; PubMed=10623306;
RA Schwemle M., Jehle C., Formella S., Staeheli P.;
RT "Sequence similarities between human Bornavirus isolates and
RT laboratory strains question their human origin.";
RL Lancet 354:1973-1974 (1999).
DR EMBL; AJ250179; CAB87242.1; -
FT NON_TER 39 39
FT NON_TER 39 39
SQ SEQUENCE 39 AA; 4205 MW; D5A828ED9C2EF134 CRC64;

Query Match
Best Local Similarity 100.0%; Score 21; DB 12; Length 39;
Pred. No. 2.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
DB 4 HARL 7

RESULT 4
O9S083 PRELIMINARY; PRT; 42 AA.
ID O9S083

```

```

AC O9S083;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, last annotation update)
DE R2R3-MYB protein (Fragment).
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OC NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-CV. LANDSBERG ERECTA; TISSUE=FLOWER;
RX Rabinowicz P.D., Roberts C., Grotewold E.;
RT "Novel Arabidopsis R2R3 Myb genes expressed in the flower.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 1 MYB-LIKE DOMAIN.
DR EMBL; AF101050; AAF00581.1; -
DR HSSP; P01103; IPOM.
DR InterPro; IPR001005; MYB-DNA-binding.
DR Pfam; PF00249; myb-DNA-binding; 1.
DR SMART; SM00395; SANT; 1.
DR PROSITE; PS50090; MYB_3; 1.
KW DNA-binding; Nuclear protein.
FT NON_TER 1 1
FT NON_TER 42 42
SQ SEQUENCE 42 AA; 4919 MW; 4CB2418E68408816 CRC64;

Query Match
Best Local Similarity 100.0%; Score 21; DB 10; Length 42;
Pred. No. 2.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
DB 25 HARL 28

RESULT 5
O8U546 PRELIMINARY; PRT; 44 AA.
ID O8U546
AC O8U546;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, last annotation update)
DE AGR_C.4298P.
GN AGR_C.4298P.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OC NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=21608551; PubMed=11743194;
RX Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M., Mullin L.,
RA Gurallo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Liu F.,
RA Houttel K., Gordon J., Vaudin M., Ischouk O., Epp A., Liu F.,
RA Wollam C., Allinger M., Doughy D., Scott C., Lappas C., Martelz B.,
RA Flanagan C., Crowell C., Gurson J., Lomo C., Seear C., Strub G.,
RA Cielo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens C58.";
RL Science 294:2323-2328 (2001).
DR EMBL; AE008151; AAK88107.1; -
SQ SEQUENCE 44 AA; 4868 MW; B29035AD358B6A8C CRC64;

Query Match
Best Local Similarity 100.0%; Score 21; DB 16; Length 44;
Pred. No. 2.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
DB 12 HARL 15

```

RESULT 6
 08Y1N1 PRELIMINARY: PRT: 45 AA.
 AC 08Y1N1: PRELIMINARY: PRT: 45 AA.
 DT 01-MAR-2002 (TREMBlrel. 20, Created)
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Putative pilin protein.
 GN RSC0659 OR RS01563.
 OS Ralstonia solanacearum (Pseudomonas solanacearum).
 OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
 OC Ralstonia.
 OX NCBI_TaxID=305;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-GM11000;
 RX MEDLINE=21681879; PubMed=11823852;
 RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
 RA Ariat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
 RA Chandler M., Choise N., Claudel-Renard C., Cunnac S., Demange N.,
 RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,
 RA Sigulier P., Thebaud P., Boucher C.A.,
 RA Weissenbach J., Whalen M., Wincker P., Levy M.,
 RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
 RL Nature 415:497-502(2002).
 DR EMBL: AL646060; CAD14189.1; -
 KW Complete proteome.
 SQ SEQUENCE 45 AA; 4776 MW; 9C76F7F9D65B3F05 CRC64;

Query Match 100.0%; Score 21; DB 16; Length 45;
 Best Local Similarity 100.0%; Pred. No. 2.7e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 HARL 4
 DB 12 HARL 15

RESULT 7
 097152 PRELIMINARY: PRT: 52 AA.
 AC 097152: PRELIMINARY: PRT: 52 AA.
 DT 01-MAY-1999 (TREMBlrel. 10, Created)
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE Reverse transcriptase (Fragment).
 GN RI-ELEMENT\POL.
 OS Drosophila arizonae (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Preygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Phnydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7263;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Tovar F.J., Rodarte S.R., Almeida R.C., Leoncini O.,
 RT "RI element in Drosophila arizonae.";
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF114256; AAD18018.1; -
 DR FlyBase: FBgn0043769; Dair\RI-element\pol.
 KW RNA-directed DNA polymerase.
 FT NON_TER 1 52
 FT NON_TER 1 52
 SQ SEQUENCE 52 AA; 6215 MW; B511A4FC41B2A40A CRC64;

Query Match 100.0%; Score 21; DB 5; Length 52;
 Best Local Similarity 100.0%; Pred. No. 3.1e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 HARL 4
 DB 26 HARL 29

RESULT 8
 08XUV9 PRELIMINARY: PRT: 57 AA.
 AC 08XUV9: PRELIMINARY: PRT: 57 AA.
 DT 01-MAR-2002 (TREMBlrel. 20, Created)
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Hypothetical protein RSC3074.
 GN RSC3074 OR RS00523.
 OS Ralstonia solanacearum (Pseudomonas solanacearum).
 OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
 OC Ralstonia.
 OX NCBI_TaxID=305;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-GM11000;
 RX MEDLINE=21681879; PubMed=11823852;
 RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
 RA Ariat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
 RA Chandler M., Choise N., Claudel-Renard C., Cunnac S., Demange N.,
 RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,
 RA Sigulier P., Thebaud P., Boucher C.A.,
 RA Weissenbach J., Whalen M., Wincker P., Levy M.,
 RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
 RL Nature 415:497-502(2002).
 DR EMBL: AL646073; CAD16783.1; -
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 57 AA; 6465 MW; 034B74EB142BEDD2 CRC64;

Query Match 100.0%; Score 21; DB 16; Length 57;
 Best Local Similarity 100.0%; Pred. No. 3.4e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 HARL 4
 DB 39 HARL 42

RESULT 9
 027782 PRELIMINARY: PRT: 59 AA.
 AC 027782: PRELIMINARY: PRT: 59 AA.
 DT 01-JAN-1998 (TREMBlrel. 05, Created)
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Hypothetical protein MTH1750.
 GN MTH1750.
 OS Methanobacterium thermoautotrophicum.
 OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
 OC Methanobacteriacea; Methanobacter.
 OX NCBI_TaxID=187420;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
 RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
 RA Harrison D., Hoang L., Keagle P., Lunn W., Pothier B., Qiu D.,
 RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
 RA Jivani N., Caruso A., Bush D., Sater H., Patwell J., Frabaker S.,
 RA McDougall S., Shimer G., Goyal A., Pietrovski S., Church G.M.,
 RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
 RT "Complete genome sequence of Methanobacterium thermoautotrophicum
 deltaH: functional analysis and comparative genomics.";
 RL J. Bacteriol. 179:7135-7155(1997).
 DR EMBL: AE000930; AAB86220.1; -
 DR InterPro: IPR004820; Cytidylyltransf.
 DR Pfam: PF01467; Cytidylyltransf. 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 59 AA; 6658 MW; 6559DB837EED72F0 CRC64;

Query Match 100.0%; Score 21; DB 17; Length 59;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
DB 15 HARL 18

RESULT 10

ID 007026 PRELIMINARY; PRT; 65 AA.

AC 007026; 01-NOV-1996 (TREMBLREL. 01, Created)
DT 01-NOV-1996 (TREMBLREL. 01, Last sequence update)
DE 01-NOV-1998 (TREMBLREL. 08, Last annotation update)
GN RFBQ protein.

OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;

RP SEQUENCE FROM N.A.
RC STRAIN-017 / BIOTYPE EL TOR / SEROTYPE OGAWA;
RX MEDLINE-92212870; PubMed-1372980;

RA Stroehrer U.H., Karageorgos L.E., Morona R., Manning P.A.;
RT "Serotype conversion in Vibrio cholerae O1.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:2566-2570(1992).

RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-017 / BIOTYPE EL TOR / SEROTYPE OGAWA;
RX MEDLINE-95309704; PubMed-7540582;

RA Manning P.A., Stroehrer U.H., Karageorgos L.E., Morona R.;
RT "Putative O-antigen transport genes within the rfb region of Vibrio
cholerae O1 are homologous to those for capsule transport.";

DR EMBL; X59554; CAA42147.1;
SQ SEQUENCE 65 AA; 7577 MW; 59FA6094E686E544 CRC64;

Query Match 100.0%; Score 21; DB 2; Length 65;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
DB 53 HARL 56

RESULT 11

ID 08UJ23 PRELIMINARY; PRT; 67 AA.

AC 08UJ23; 01-JUN-2002 (TREMBLREL. 21, Created)
DT 01-JUN-2002 (TREMBLREL. 21, Last sequence update)
DE 01-JUN-2002 (TREMBLREL. 21, Last annotation update)
GN Hypothetical protein Atu0115.

OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OX NCBI_TaxID=176299;

RP SEQUENCE FROM N.A.
RC MEDLINE-21608550; PubMed-11743193;

RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RX Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F., Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Boyce D. Sr.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,

RA Kutayavir T., Levy R., Li M.-J., McClelland E., Palmeri A.,
RA Raymond C., Rouse G., Saenphitumachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,

RA Nester E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
strain C58.";
RL Science 294:2317-2323(2001).

DR EMBL; AEO08985; AAL41140.1;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 67 AA; 7608 MW; 351679E2DE08499 CRC64;

Query Match 100.0%; Score 21; DB 16; Length 67;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
DB 53 HARL 56

RESULT 12

ID 08VAB5 PRELIMINARY; PRT; 68 AA.

AC 08VAB5; 01-MAR-2002 (TREMBLREL. 20, Created)
DT 01-MAR-2002 (TREMBLREL. 20, Last sequence update)
DE 01-JUN-2002 (TREMBLREL. 21, Last annotation update)
GN WsV511 (WSSV037).

OS White spot syndrome virus (WSSV).
OC Viruses; unclassified viruses.
OX NCBI_TaxID=92652;

RP SEQUENCE FROM N.A.
RC MEDLINE-21548311; PubMed-11689662;

RA Yang F., He J., Lin X., Li Q., Pan D., Zhang X., Xu X.;
RT "Complete genome sequence of the shrimp white spot bacilliform
virus.";
RL J. Virol. 75:11811-11820(2001).

RN [2]
RP SEQUENCE FROM N.A.
RA Yang F., He J., Lin X., Li Q., Pan D., Zhang X., Xu X.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.

RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-TAIWAN;

RA MEDLINE-20517548; PubMed-11062040;
RX Tsai M.F., Yu H.T., Tzeng H.F., Leu J.H., Chou C.M., Huang C.J.,
RA Wang C.H., Lin J.Y., Kou G.H., Lo C.F.;

RT "Identification and characterization of a shrimp white spot syndrome
cellular-type thymidine kinase and thymidylate kinase.";
RL Virology 277:100-110(2000).

RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN-TAIWAN;

RA MEDLINE-21844071; PubMed-11853398;
RX Chen L.L., Leu J.H., Huang C.J., Chou C.M., Chen S.M., Wang C.H.,
RA Lo C.F., Kou G.H.;

RT "Identification of a nucleocapsid protein (VP35) gene of shrimp white
spot syndrome virus and characterization of the motif important for
targeting VP35 to the nuclei of transfected insect cells.";

RL Virology 293:44-53(2002).

RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN-TAIWAN;

RA Lo C.-F., Kou G.-H.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF440570; AAL88905.1;
SQ SEQUENCE 68 AA; 8104 MW; 61908AE8049D134D CRC64;

Query Match 100.0%; Score 21; DB 12; Length 68;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4

Db 24 HARL 27

RESULT 13

09B094 ID 09B094 PRELIMINARY; PRT; 69 AA.
 AC 09B094;
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DE Similar to hypothetical protein FLJ20378.
 OS Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OX Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 RN NCBI_TaxID=9606;
 RP STRAIN=CV. NIPPONBARE;
 RC Sasaki T., Matsumoto T., Yamamoto K.;
 RA clon:B1157F09.18;
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AP003207; BAB64092.1;
 KW EMBL; BC002811; AAH02811.1;
 SQ SEQUENCE 69 AA; 7649 MW; 48BA04F9785E8098 CRC64;

Query Match
 Best Local Similarity 100.0%; Score 21; DB 4; Length 69;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 HARL 4
 Db 61 HARL 64

RESULT 14

09Y21 ID 09Y21 PRELIMINARY; PRT; 69 AA.
 AC 09Y21;
 DT 01-MAY-1999 (TREMBlrel. 10, Created)
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
 DE Hypothetical 7.8 kDa protein.
 OS Christococcus funiformis
 OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
 OX Nucleopolyhedrovirus.
 RN NCBI_TaxID=10448;
 RP SEQUENCE FROM N.A.
 RC STRAIN=IRELAND;
 RA Polunienko A., Krell P.J.;
 RT "Identification of the ORF 8/6 gene in the EGT-TAPI intergenic region
 of a baculovirus pathogenic to the spruce budworm, Cf MNPV."
 RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U70432; AAD10302.1;
 KW Hypothetical protein.
 SQ SEQUENCE 69 AA; 7847 MW; 95595EAD798879F3 CRC64;

Query Match
 Best Local Similarity 100.0%; Score 21; DB 12; Length 69;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 HARL 4
 Db 45 HARL 48

RESULT 15

094301 ID 094301 PRELIMINARY; PRT; 70 AA.
 AC 094301;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DE Hypothetical protein
 GN RVI1684 OR MTG1125.06 OR MTG1724.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Firmicutes; Actinobacteria;

DE B1157F09.18 protein.
 GN B1157F09.18.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Euphorbiaceae; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OX NCBI_TaxID=4530;
 RN NCBI_TaxID=4530;
 RP STRAIN=CV. NIPPONBARE;
 RC Sasaki T., Matsumoto T., Yamamoto K.;
 RA clon:B1157F09.18;
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AP003207; BAB64092.1;
 KW EMBL; AP003207; BAB64092.1;
 SQ SEQUENCE 70 AA; 8053 MW; 1038D3DE5DEF6AFD CRC64;

Query Match
 Best Local Similarity 100.0%; Score 21; DB 10; Length 70;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 HARL 4
 Db 45 HARL 48

RESULT 16

09IK63 ID 09IK63 PRELIMINARY; PRT; 74 AA.
 AC 09IK63;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DE ORF1 (Fragment).
 OS TT virus.
 OC Viruses; ssDNA viruses; unclassified ssDNA viruses.
 OX NCBI_TaxID=68887;
 RN NCBI_TaxID=68887;
 RP SEQUENCE FROM N.A.
 RC MEDLINE-20216631; PubMed-10751348;
 RA Rodriguez-Inigo E., Casqueiro M., Barolome J., Ortiz-Movilla N.,
 RA Lopez-Alcorcho J.M., Herrero M., Manzanilla F., Oliva H.,
 RA Carreno V.;
 RT "Detection of TT virus DNA in liver biopsies by in situ
 hybridization."
 RL Am. J. Pathol.
 DR EMBL; AF216403; AAF79086.1;
 DR InterPro: IPR004219; TrTVirus_Unk.
 DR Pfam: PF02956; TT-ORF1; I.
 FT NON_TER
 SQ SEQUENCE 74 AA; 8318 MW; CEE6503663BFB02 CRC64;

Query Match
 Best Local Similarity 100.0%; Score 21; DB 12; Length 74;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 HARL 4
 Db 45 HARL 48

RESULT 17

033186 ID 033186 PRELIMINARY; PRT; 74 AA.
 AC 033186;
 DT 01-JAN-1998 (TREMBlrel. 05, Created)
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 DE Hypothetical protein RVI1684
 GN RVI1684 OR MTG1125.06 OR MTG1724.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Firmicutes; Actinobacteria;

Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-H37RV;
 RX MEDLINE=98295987; PubMed=9634230; Garnier T., Churcher C., Harris D., Cole S.T., Brosch R., Parkhill J., Barry C.E. III, Tekala F., Gordon S.V., Eigmeier K., Gas S., Chillingworth T., Connor R., Badcock K., Basham D., Brown D., Chillingworth T., Connor R., Davies K., Devlin K., Feldwell T., Gentles S., Hamlin N., Holroyd S., Hornby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L., Oliver S., Osborne J., Quail M.A., Raeburn M.A., Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R., Stulton J.E., Taylor K., Whitehead S., Barrell B.G.; "Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence."
 RT Nature 393:537-544(1998).
 RL [2]
 RN SEQUENCE FROM N.A.
 RP STRAIN-CDC 1551 / OSHKOSH; Eisen J.A., Carpenter L., White O., Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O., Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E., Kolony J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L., Delner A., Uterback T., Weidman J., Khouri H., Gill J., Mikula A., Bishai W.; "Whole genome comparison of Mycobacterium tuberculosis clinical and laboratory strains." to the EMBL/GenBank/DBJ databases.
 RT Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 RL EMBL: Z98268; CAB10941.1;
 DR EMBL: A6007035; AAK45992.1;
 DR TIGR: W1724;
 DR TubercuList; Rv1684;
 DR Hypothetical protein; Complete proteome.
 KW SEQUENCE 74 AA; 8289 MW; 84784E241903C263 CRC64;
 SQ
 Query Match 100.0%; Score 21; DB 16; Length 74;
 Best Local Similarity 100.0%; Pred. No. 4.3e+02; Indels 0; Gaps 0;
 Matches 4; Conservative 0; Mismatches 0;
 OY 1 HARL 4
 Db 61 HARL 64
 RESULT 18
 ID 065817 PRELIMINARY; PRT: 75 AA.
 AC 065817;
 DT 01-AUG-1998 (TREMBlrel. 07, Created)
 DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Chlorophyll a/b-binding protein (Fragment).
 GN CP29.
 OS Hordeum vulgare (Barley).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Eukaryota; Viridiplantae; Magnoliophyta; Liliopsida; Poales; Poaceae; Poaceae;
 CC Trilicaceae; Hordeum.
 OC NCBI_TaxID=4513;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. CARINA. TISSUE-LEAF;
 RX Humbeck K., Krupinska K.; "Differential regulation of chlorophyll a/b-binding proteins and ELIP in mature and senescencing leaves of field grown barley."
 RT Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
 RL EMBL: AJ006296; CA06961.1;
 DR InterPro: IPR001344; Chloro_Abbind.
 DR Pfam: PF00504; Chloro_a-b-bind. 1.
 DR ProDom: PD000275; Chloro_Abbind; 1.
 FT NON TER 1
 RA SEQUENCE 75 AA; 7988 MW; 5442D30B9E61057D CRC64;
 SQ
 Query Match 100.0%; Score 21; DB 10; Length 75;

Best Local Similarity 100.0%; Pred. No. 4.3e+02; Indels 0; Gaps 0;
 Matches 4; Conservative 0; Mismatches 0;
 OY 1 HARL 4
 Db 30 HARL 33
 RESULT 19
 ID 09XB07 PRELIMINARY; PRT: 79 AA.
 AC 09XB07;
 DT 01-NOV-1999 (TREMBlrel. 12, Created)
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Acyl carrier protein (ACP).
 GN TAB.
 OS Myxococcus xanthus.
 CC Bacteria; Proteobacteria; delta subdivision; Myxobacteria;
 CC Myxococcales; Cystobacterineae; Myxococcaceae; Myxococcus.
 OC NCBI_TaxID=34;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ER-15;
 RX MEDLINE=20055630; PubMed=10589713;
 RA Paltan Y., Orr E., Ron E.Z., Rosenberg E.; "Genetic and functional analysis of genes required for the post-modification of the polyketide antibiotic TA of Myxococcus xanthus."
 RT Microbiology 145:3059-3067(1999).
 RL EMBL: AJ132503; CAB46501.1;
 DR InterPro: IPR003880; ACP_DOMAIN; 1.
 DR PROSITE: PS50075; ACP_DOMAIN; 1.
 KW Phosphopantetheine.
 SQ SEQUENCE 79 AA; 8730 MW; 4F1C5F67B8192F7 CRC64;
 Query Match 100.0%; Score 21; DB 2; Length 79;
 Best Local Similarity 100.0%; Pred. No. 4.6e+02; Indels 0; Gaps 0;
 Matches 4; Conservative 0; Mismatches 0;
 OY 1 HARL 4
 Db 76 HARL 79
 RESULT 20
 ID 09PDG2 PRELIMINARY; PRT: 80 AA.
 AC 09PDG2;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Hypothetical protein Xf1417.
 GN Xf1417.
 OS Xylella fastidiosa.
 CC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
 CC Xylella.
 OC NCBI_TaxID=2371;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-9A5C;
 RX MEDLINE=20365717; PubMed=10910347;
 RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Agencio M., Alvarenga R., Alves L.M.C., Araya J.E., Bala G.S., Brites M.R.S., Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Brites M.R.S., Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrier H., Coutinho L.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M., Coutinho L.B., Cristofani M., Dias-Neto E., Docena C., El-Dorri H., Facinanci A.P., Ferreira A.J.S., Ferreira V.C.A., Ferris J.A., Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furian L.R., Fraga J.S., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A., Garnier M., Goldman J.D., Junqueira M.L., Kemper E.L., Kiteima J.P., Ho P.L., Hohnselt J.D., Junqueira M.L., Kempf E.L., Leite L.C.C., Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Machado J.A., Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,

RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
 RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
 RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
 RA Moon D.H., Nagel M.A., Nascimento A.L.T.O., Netto L.E.S.,
 RA Nham A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
 RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
 RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
 RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
 RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
 RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
 RA da Silveira J.F., Silvestri M.L.Z., Silveira W.J., de Souza A.A.,
 RA de Souza A.P., Terezi M.F., Truffi D., Tsai S.M., Tuhako M.H.,
 RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
 RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.,
 RT "The genome sequence of the plant pathogen *Xylella fastidiosa*,"
 RL Nature 406:151-159(2000).
 DR EMBL: AE003972; AAF84226.1; -
 KM Hypothetical protein; Complete proteome.
 SQ SEQUENCE 80 AA; 8778 MW; 93765FBA2D96C7D CRC64;

Query Match 100.0%; Score 21; DB 16; Length 80;
 Best Local Similarity 100.0%; Pred. No. 4.6e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
 ||||
 DB 53 HARL 56

RESULT 21

ID 08XY77 PRELIMINARY; PRT; 80 AA.
 AC 08XY77;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Probable transmembrane protein.
 GN RSC2688 OR RS00033.
 OS Ralstonia solanacearum (Pseudomonas solanacearum).
 OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
 OC Ralstonia.
 OX NCBI_TaxID=305;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GM1100;
 RX MEDLINE=2161879; PubMed=11823852;
 RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
 RA Arlat M., Billault A., Brothier P., Camus J.C., Catholico L.,
 RA Chandler M., Choisme N., Claudel-Renard C., Cunac S., Demange N.,
 RA Gaspin C., Lave M., Moisan A., Robert C., Saurin W., Schlex T.,
 RA Sigulier P., Thebaud P., Whalen M., Wincker P., Levy M.,
 RA Weissenbach J., Boucher C.A.,
 RT "Genome sequence of the plant pathogen *Ralstonia solanacearum*,"
 RL Nature 415:497-502(2002).
 DR EMBL: AF646071; CAD16395.1; -
 KM Complete proteome.
 SQ SEQUENCE 80 AA; 8473 MW; D6373F7D6B0CB69 CRC64;

Query Match 100.0%; Score 21; DB 16; Length 80;
 Best Local Similarity 100.0%; Pred. No. 4.6e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
 ||||
 DB 46 HARL 49

RESULT 22
 09M675
 ID 09M675 PRELIMINARY; PRT; 81 AA.
 AC 09M675;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)

DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
 DE Cystathionine-gamma synthase (Fragment).
 OS Cucumis melo (Muskmelon).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Euphorbia; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eucosids I; Cucurbitales; Cucurbitaceae; Cucumis.
 OX NCBI_TaxID=3656;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=20177879; PubMed=10712562;
 RX Hadfield K.A., Dang T., Guis M., Pech J.-C., Bouzayen M.,
 RA Bennett A.B.,
 RT "Characterization of ripening-regulated cDNAs and their expression in
 ethylene-suppressed *Charantia melon* fruit.";
 RL Plant Physiol. 122:977-983(2000).
 DR EMBL: AF206626; AAF64422.1; -
 FT NON_TER 1 1
 FT 81 81
 SQ SEQUENCE 81 AA; 8883 MW; 5B043AA22FD8D247 CRC64;

Query Match 100.0%; Score 21; DB 10; Length 81;
 Best Local Similarity 100.0%; Pred. No. 4.7e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
 ||||
 DB 6 HARL 9

RESULT 23

ID 09WGD5 PRELIMINARY; PRT; 83 AA.
 AC 09WGD5;
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE E5 protein.
 GN E5.
 OS Human papillomavirus type 16.
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
 OC Papillomavirus.
 OX NCBI_TaxID=10581;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=IS347;
 RX MEDLINE=96079021; PubMed=7494284;
 RA Yamada T., Wheeler C.M., Halpern A.L., Stewart A.C., Hildesheim A.,
 RA Jenison S.A.,
 RT "Human papillomavirus type 16 variant lineages in United States
 RT populations characterized by nucleotide sequence analysis of the E6,
 RT L2, and L1 coding segments.";
 RL J. Virol. 69:7743-7753(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=IS347;
 RX MEDLINE=99190418; PubMed=10091997;
 RA Eriksson A., Herron J.R., Yamada T., Wheeler C.M.,
 RT "Human papillomavirus type 16 variant lineages characterized by
 RT nucleotide sequence analysis of the E5 coding segment and the E2 hinge
 RT region.";
 RL J. Gen. Virol. 80:595-600(1999).
 DR EMBL: AF120704; AAD24032.1; -
 DR InterPro: IPR004270; Papilloma-E5.
 DR Pfam: PF03025; Papilloma-E5; 1.
 SQ SEQUENCE 83 AA; 9365 MW; 40612DC9973C38BD CRC64;

Query Match 100.0%; Score 21; DB 12; Length 83;
 Best Local Similarity 100.0%; Pred. No. 4.8e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
 ||||
 DB 77 HARL 80

RESULT 24

09K5X7 PRELIMINARY; PRT; 83 AA.
AC 09K5X7;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Hypothetical protein BH3939.
GN BH3939.
OS Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Bacillaceae; Bacillus.
OX NCBI_TaxID=86665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-C-125 / JCM 9153;
RX MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hirama C., Nakamura T., Ogasawara N., Kihara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis."
RL Nucleic Acids Res. 28:4317-4331(2000).
DR EMBL; AP001520; BAB07678.1;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 83 AA; 9977 MW; 1EE8407781E52C20 CRC64;

Query Match 100.0%; Score 21; DB 16; Length 83;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
DB 35 HARL 38

RESULT 25

066074 PRELIMINARY; PRT; 85 AA.
AC 066074;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Virion protein (Fragment).
OS Canine herpesvirus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae.
OX NCBI_TaxID=371110;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-MILOU;
RX MEDLINE=96145133; PubMed=8558127;
RA Remond M., Sheldrick P., Lebreton F., Nardeux P., Foulon T.;
RT "Gene organization in the UL region and inverted repeats of the canine
RT herpesvirus genome."
RL J. Gen. Virol. 77:37-48(1996).
DR EMBL; X90444; CA62069.1;
DR InterPro: IPR002660; Herpes_UL6.
DR Pfam; PF01763; Herpes_UL6; 1.
DR Prodom; PD003210; Herpes_UL6; 1.
FT NON_TER
FT NON_TER
GN NON_TER
SQ SEQUENCE 85 AA; 9584 MW; 3DF46D8AAE18ADCF CRC64;

Query Match 100.0%; Score 21; DB 12; Length 85;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
DB 30 HARL 33

RESULT 26

08UIC1 PRELIMINARY; PRT; 85 AA.
AC 08UIC1;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Hypothetical protein Atu0376.
GN Atu0376 OR AGR_C_659.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608550; PubMed=11743193;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitaajima J.P.,
RA Okura Y.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kutayavin T., Levy R., Li M.-J., McClelland E., Palmeri A.,
RA Raymond C., Rouse G., Saepthimachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nester E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT C58."
RL Science 294:2317-2323(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608551; PubMed=11743194;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Houriello B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
RA Hounell K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
RA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Marleiz B.,
RA Flanagan C., Crowell C., Gursun J., Lomo C., Sear C., Strud G.,
RA Cielo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens C58."
RL Science 294:2323-2328(2001).
DR EMBL; AE009008; AAL4398.1;
DR EMBL; AE007975; AAK86193.1;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 85 AA; 9466 MW; 518B33B66A33219 CRC64;

Query Match 100.0%; Score 21; DB 16; Length 85;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
DB 7 HARL 10

RESULT 27

09FD17 PRELIMINARY; PRT; 87 AA.
AC 09FD17;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Acyl carrier protein.
GN RUBC

OS Streptomyces collinus.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomycetes.
OX NCBI_TaxID=42684;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-DSM2012;

RA Martin R., Bailey J.E., Minas W.,
RT "Rudromycin polyketide synthase from Streptomyces collinus DSM2012.";
RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF293355; AAC03069.1; -
DR HSSP: 002054; 2AP8.
DR InterPro: IPR003880; Pantine_attach.
DR Pfam: PF00550; pp-binding; 1.
DR PROSITE: PS50075; ACP_DOMAIN; 1.
DR PROSITE: PS00012; PHOSPHOPANTETHEINE; UNKNOWN_1.
KW Phosphopantetheine.
SQ SEQUENCE 87 AA; 9792 MW; 235BDCC875F14C0 CRC64;

Query Match 100.0%; Score 21; DB 2; Length 87;
Best Local Similarity 100.0%; Pred. No. 5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
DB 79 HARL 82

RESULT 28

O9FBZ3 PRELIMINARY; PRT; 87 AA.
AC O9FBZ3;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, last annotation update)
DE Hypothetical protein SC07189.
GN SC07189 OR SC8A11.17C.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2) / M145;
RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitz E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Mletzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL: AL391041; CAC01589.1; -
KW Hypothetical protein.
SQ SEQUENCE 87 AA; 9955 MW; F0008D1049B11D4B CRC64;

Query Match 100.0%; Score 21; DB 16; Length 87;
Best Local Similarity 100.0%; Pred. No. 5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
DB 41 HARL 44

RESULT 29

O9YKD5 PRELIMINARY; PRT; 88 AA.
AC O9YKD5;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE Picoeal protein.
GN V2.
OS tomato yellow leaf curl virus - II.
OC Viruses; ssDNA viruses; Geminitoviridae; Begomovirus.

OX NCBI_TaxID=83229;
RN [1]
RP SEQUENCE FROM N.A.
RA Peterschmitt M., Granier M., Aboulama S.;
RT "First report of tomato yellow leaf curl geminivirus virus in
RT Morocco.";
RL Plant Dis. 83:1074-1074(1999).
DR EMBL: AJ133491; CAB38575.1; -
DR InterPro: IPR002511; Gemin1_V1.
DR Pfam: PF01524; Gemin1_V1; 1.
DR ProDom: PD002978; Gemin1_V1; 1.
KW Coat protein.
SQ SEQUENCE 88 AA; 10378 MW; 32D149F7EB0F9C39 CRC64;

Query Match 100.0%; Score 21; DB 12; Length 88;
Best Local Similarity 100.0%; Pred. No. 5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
DB 65 HARL 68

RESULT 30

O93SB6 PRELIMINARY; PRT; 92 AA.
AC O93SB6;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE Hypothetical 10.1 kDa protein.
GN Frankia sp. Ar13.
OS Plasmid pF031.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Frankineae; Frankiaceae; Frankia.
OX NCBI_TaxID=1858;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-AR13;
RX MEDLINE-21184415; PubMed-11287155;
RA Lavigne C., Louis D., Perriere G., Briolay J., Normand P.,
RA Courvoyer B.;
RT "Analysis of pF031, a 8551-bp cryptic plasmid from the symbiotic
RT nitrogen-fixing actinomycete Frankia.";
RL FEMS Microbiol. Lett. 197:111-116(2001).
DR EMBL: AJ297945; CAC39346.1; -
KW Hypothetical protein; Plasmid.
SQ SEQUENCE 92 AA; 10054 MW; B6A0D4F13CF6FED8 CRC64;

Query Match 100.0%; Score 21; DB 2; Length 92;
Best Local Similarity 100.0%; Pred. No. 5.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
DB 24 HARL 27

RESULT 31

O18842 PRELIMINARY; PRT; 93 AA.
AC O18842;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-JAN-1998 (TREMBLrel. 05, last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE C54G10.1 protein.
GN C54G10.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.

RA Matthews L.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-99069613; PubMed-9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL: Z75332; CAA9809.1; -.
SQ SEQUENCE 93 AA; 10885 MW; C00F5DC1515DE15B CRC64;

Query Match 100.0%; Score 21; DB 5; Length 93;
Best Local Similarity 100.0%; Pred. No. 5.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HARL 4
DB 19 HARL 22

RESULT 32
Q91113 PRELIMINARY; PRT; 93 AA.
ID 091113
AC 091113
DT 01-MAR-2001 (TRENBLREL. 16, Created)
DT 01-MAR-2001 (TRENBLREL. 16, Last sequence update)
DT 01-OCT-2001 (TRENBLREL. 18, Last annotation update)
DE Hypothetical protein PA2292.
GN PA2292.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
RX [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 15692 / PA01;
RX MEDLINE-20437337; PubMed-10984043;
RA Stever C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
Garber R.L., Goltzy L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
Brody L.L., Coulter S.N., Folger K.R., Kas A., Iatrig K., Lam R.M.,
Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
Reiter J., Seiler M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
opportunistic pathogen.";
RL Nature 406:959-964(2000).
DR EMBL: AE004655; AAC05680.1; -.
DR InterPro: IPR000886; ER_target.
DR PROSITE: PS00014; ER_TARGET; UNKNOWN_1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 93 AA; 10732 MW; A38541EC2F12AF22 CRC64;

Query Match 100.0%; Score 21; DB 16; Length 93;
Best Local Similarity 100.0%; Pred. No. 5.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HARL 4
DB 31 HARL 34

RESULT 33
O8ZK19 PRELIMINARY; PRT; 95 AA.
ID 08ZK19
AC 08ZK19
DT 01-MAR-2002 (TRENBLREL. 20, Created)
DT 01-MAR-2002 (TRENBLREL. 20, Last sequence update)
DT 01-MAR-2002 (TRENBLREL. 20, Last annotation update)
DE Putative integrase.
GN STM4488.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

OC Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE-21534948; PubMed-11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
LT2.";
RL Nature 413:852-856(2001).
DR EMBL: AE008910; AAL23306.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 95 AA; 10938 MW; 47F9372FCED1FED CRC64;

Query Match 100.0%; Score 21; DB 16; Length 95;
Best Local Similarity 100.0%; Pred. No. 5.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HARL 4
DB 10 HARL 13

RESULT 34
Q38025 PRELIMINARY; PRT; 96 AA.
ID 038025
AC 038025
DT 01-NOV-1996 (TRENBLREL. 01, Created)
DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)
DT 01-DEC-2001 (TRENBLREL. 19, Last annotation update)
DE ORF 4.
GN 4.
OS Bacteriophage phi-C31.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;
OC Lambda-like viruses.
OX NCBI_TaxID=10719;
RX [1]
RP SEQUENCE FROM N.A.
RC MEDLINE-94374705; PubMed-8088546;
RA Hartley N.M., Murphy G.O., Bruton C.J., Chater K.F.;
RT "Sequence of the essential early region of phi C31, a temperate phage
of Streptomyces spp. with unusual features in its lytic development.";
RL gene 147:29-40(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-NORWICH STOCK;
RA Smith M.C.M.;
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-NORWICH STOCK;
RX MEDLINE-99162580; PubMed-10051617;
RA Hendrix R.W., Smith M.C.M., Burns N., Ford M.E., Hatfull G.F.;
RT "Evolutionary relationships among diverse bacteriophages and
RT prophages: all the world's a phage.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:2192-2197(1999).
DR EMBL: X76288; CAA53915.1; -.
DR EMBL: AJ006589; CAA07129.1; -.
SQ SEQUENCE 96 AA; 10601 MW; 6C136673B5B4E63 CRC64;

Query Match 100.0%; Score 21; DB 9; Length 96;
Best Local Similarity 100.0%; Pred. No. 5.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HARL 4
DB 82 HARL 85

RESULT 35

09M9W0 PRELIMINARY; PRT; 97 AA.
 ID 09M9W0
 AC 09M9W0
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE F18C1.18 protein.
 GN F18C1.18
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC Eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Lin X., Kaul S., Town C.D., Beito M., Creasy T.H., Haas B., Wu D.,
 RA Rongling C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,
 RA Bowman C.L., White O., Niernan W.C., Fraser C.M.;
 RT "Arabidopsis thaliana chromosome III BAC F18C1 genomic sequence."
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC011620; AAF26142.1;
 SQ SEQUENCE 97 AA; 10967 MW; 3C45FBC676E27F9 CRC64;

Query Match 100.0%; Score 21; DB 10; Length 97;
 Best Local Similarity 100.0%; Pred. No. 5.5e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
 ||||
 DB 52 HARL 55

RESULT 36

087620 PRELIMINARY; PRT; 99 AA.
 ID 087620
 AC 087620
 DT 01-NOV-1998 (TREMBLrel. 08, Created)
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE O-halobenzoate dioxygenase ferredoxin.
 GN OHBD.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas
 OX NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=JB2;
 RX MEDLINE=21455015; PubMed=11571162;
 RA Hickey W.J., Sabat G., Turfot A.S., Ament A.R., Perez-Leisher J.;
 RT "Cloning, Nucleotide Sequencing, and Functional Analysis of a Novel,
 RT Mobile Cluster of Biodegradation Genes from Pseudomonas aeruginosa
 RT Strain JB2."
 RL Appl. Environ. Microbiol. 67:4603-4609(2001).
 DR EMBL: AF087482; AAC69486.1;
 DR HSSP; P37332; IFO7.
 DR InterPro; IPR001281; R1eske.
 DR Pfam; PF00355; R1eske; 1.
 KW Dioxygenase.
 SQ SEQUENCE 99 AA; 10836 MW; 07C198B76AF24EA3 CRC64;

Query Match 100.0%; Score 21; DB 2; Length 99;
 Best Local Similarity 100.0%; Pred. No. 5.6e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
 ||||
 DB 43 HARL 46

RESULT 37

092G07 PRELIMINARY; PRT; 100 AA.
 ID 092G07
 AC 092G07
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Hypothetical protein RC1065.
 GN RC1065.
 OS Rickettsia conorii.
 OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
 OC Rickettsiaceae; Rickettsiae; Rickettsia.
 OX NCBI_TaxID=781;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MALISH 7;
 RX MEDLINE=21442074; PubMed=11557893;
 RA Ogata H., Audic S., Renesto-Audiffren P., Fournier P.-E., Barbe V.,
 RA Samson D., Roux V., Cossart P., Weissenbach J., Claverie J.-M.,
 RA Raoult D.;
 RT "Mechanisms of evolution in Rickettsia conorii and R. prowazekii."
 RL Science 293:2093-2098(2001).
 DR EMBL: AE008657; AL03603.1;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 100 AA; 11611 MW; D4354112ACDAE167 CRC64;

Query Match 100.0%; Score 21; DB 16; Length 100;
 Best Local Similarity 100.0%; Pred. No. 5.6e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
 ||||
 DB 77 HARL 80

RESULT 38

093017 PRELIMINARY; PRT; 101 AA.
 ID 093017
 AC 093017
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE RECA (Fragment).
 GN RECA.
 OS Bacteroides ovatus.
 OC Bacteria; CRB group; Bacteroidetes; Bacteroidales; Bacteroidaceae;
 OC Bacteroides.
 OX NCBI_TaxID=28116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=F120C;
 RA Gutacker M.W., Piffaretti J.-C.;
 RT "Sequence diversity within the recA and glnA genes of B. fragilis
 RT delineates two subgroups associated to different antibiotic resistance
 RT genotypes (cepa and cfla)."
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF280747; AAK4950.1;
 DR InterPro; IPR001553; RECA.
 DR Pfam; PF00154; recA; 1.
 DR ProDom; PD000229; RECA; 1.
 DR PROSITE; PS50162; RECA_2; 1.
 FT NON_TER 1
 FT NON_TER 1
 SQ SEQUENCE 101 AA; 11016 MW; BA418B0AD44211C2 CRC64;

Query Match 100.0%; Score 21; DB 2; Length 101;
 Best Local Similarity 100.0%; Pred. No. 5.7e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
 ||||
 DB 74 HARL 77

```

RESULT 39
09JRG6 PRELIMINARY: PRT: 101 AA.
ID 09JRG6
AC 09JRG6;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Hypothetical protein NMA0377.
GN NMA0377 OR NME2058.
OS Neisseria meningitidis (serogroup A), and
OS Neisseria meningitidis (serogroup B).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=65699, 491;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=Z2491 / SEROGROUP A / SEROTYPE 4A;
RX MEDLINE=20222556; PubMed=10761919;
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holtroyd S.,
RA Jagels K., Leather S., Moule S., Mungall K., Quail M.A.,
RA Rajandream M.A., Rutherford K.M., Stimpson M., Skelton J.,
RA Whitehead S., Spratt B.G., Barrall B.G.;
RT "Complete DNA sequence of a serogroup A strain of Neisseria
RT meningitidis Z2491."
RL Nature 404:502-506(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MC58 / SEROGROUP B;
RX MEDLINE=20175755; PubMed=10710307;
RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
RA Hatt D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA Mason T., Ciecko A., Parksey D.S., Blair E., Clifton H., Clark E.B.,
RA Cotton M.D., Utterback T.R., Khouiri H., Qin H., Vamathevan J.,
RA Gill J., Scarlato V., Maignani V., Pizzi M., Grandi G., Sun L.,
RA Smith H.O., Fraser C.M., Moxon E.R., Rappunli R., Venter J.C.;
RT "Complete genome sequence of Neisseria meningitidis serogroup B strain
RT MC58."
RL Science 287:1809-1815(2000).
DR EMBL: AL162753; CAB83678.1; -
DR EMBL: AF002555; AAF42378.1; -
DR TIGR: NME2058; -
RW Hypothetical protein; Complete proteome.
SQ SEQUENCE 101 AA; 10908 MW; 18938C359C04FCB6 CRC64;

Query Match 100.0%; Score 21; DB 16; Length 101;
Best Local Similarity 100.0%; Pred. No. 5.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
Db 12 HARL 15

RESULT 40
037880 PRELIMINARY: PRT: 102 AA.
ID 037880
AC 037880;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE V1 protein (Fragment).
GN V1.
OS Tomato yellow leaf curl virus - Israel.
OC Viruses; ssDNA viruses; Gemintviridae; Begomovirus.
OX NCBI_TaxID=66366;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPAIN SP697;
RA Navas-Castillo J., Sanchez-Campos S., Diaz J.A., Saez-Alonso E.,
RA Moriones E.;

```

```

RT "First report of tomato yellow leaf curl virus-is in Spain:
RT coexistence of two different gemintviruses in the same epidemic
RT outbreak."
RL Plant Dis. 0:0-0(1997).
DR EMBL: AF022219; AAB84395.1; -
DR InterPro: IPR002511; Gemini_V1.
DR InterPro: IPR005159; WCCH.
DR Pfam: PF01524; Gemini_V1. 1.
DR Pfam: PF03716; WCCH; 1.
DR ProDom: PD002978; Gemini_V1. 1.
FT NON_TER 1
FT NON_TER 102
SQ SEQUENCE 102 AA; 11807 MW; 882P967851E67AC4 CRC64;

Query Match 100.0%; Score 21; DB 12; Length 102;
Best Local Similarity 100.0%; Pred. No. 5.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
Db 57 HARL 60

RESULT 41
09W838 PRELIMINARY: PRT: 103 AA.
ID 09W838
AC 09W838;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE K8.2.
GN K1E-1.
OS Kaposi's sarcoma-associated herpesvirus (KSHV) (Human herpesvirus 8).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Rhadinovirus.
OX NCBI_TaxID=37296;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BC-1;
RX MEDLINE=99292848; PubMed=10364304;
RA Zhu F.X., Cusano T., Yuan Y.;
RT "Identification of the immediate-early transcripts of Kaposi's
RT sarcoma-associated herpesvirus."
RL J. Virol. 73:5556-5567(1999).
DR EMBL: AF091350; AAD25323.1; -
DR EMBL: AF091348; AAD25317.1; -
DR EMBL: AF091349; AAD25320.1; -
DR InterPro: IPR001832; Claudin.
DR ProSITE: PS01346; CLAUDIN; UNKNOWN_1.
SQ SEQUENCE 103 AA; 11979 MW; A5117E9BEA4FF229 CRC64;

Query Match 100.0%; Score 21; DB 12; Length 103;
Best Local Similarity 100.0%; Pred. No. 5.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
Db 13 HARL 16

RESULT 42
08UGL9 PRELIMINARY: PRT: 103 AA.
ID 08UGL9
AC 08UGL9;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Hypothetical protein Atu01018.
GN Atu01018 OR AGR_C_1876.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=176299;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-21608550; PubMed-11743193;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura Y., Zhou Y., Chen L., Wood G.E., Almeida N.F., Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Boye D. Sr.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kutyavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,
RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nester E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT C58."
RL Science 294:2317-2323(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-21608551; PubMed-11743194;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Ouello B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
RA Houmel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
RA Wollam C., Allinger M., Doughy D., Scott C., Lappas C., Markelz B.,
RA Flanagan C., Crowell C., Gursun J., Lomo C., Sear C., Strub G.,
RA Cleio C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens C58."
RL Science 294:2323-2328(2001).
DR EMBL; AE009066; AAL42031.1; -
DR EMBL; AE008032; AAK6826.1; -
KW Hypothetical protein; Complete proteome.
SO SEQUENCE 103 AA; 11758 MW; C7F5B067D45662A CRC64;

Query Match
Best Local Similarity 100.0%; Score 21; DB 16; Length 103;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HARL 4
DB 61 HARL 64

RESULT 43
Q8U5L6 PRELIMINARY; PRT; 103 AA.
ID Q8U5L6;
AC Q8U5L6;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE AGR_C_774P.
GN AGR_C_774.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-21608551; PubMed-11743194;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Ouello B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
RA Houmel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
RA Wollam C., Allinger M., Doughy D., Scott C., Lappas C., Markelz B.,
RA Flanagan C., Crowell C., Gursun J., Lomo C., Sear C., Strub G.,
RA Cleio C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens C58."
RL Science 294:2323-2328(2001).
DR EMBL; AE007981; AAK86253.1; -
SO SEQUENCE 103 AA; 11664 MW; 010BC2E7A865CAE7 CRC64;

Query Match
Best Local Similarity 100.0%; Score 21; DB 16; Length 103;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 HARL 4
DB 45 HARL 48

RESULT 44
Q8Y0F6 PRELIMINARY; PRT; 103 AA.
ID Q8Y0F6;
AC Q8Y0F6;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Probable ferredoxin subunit of A RING-hydroxylating dioxygenase
DE oxidoreductase protein (EC 1.-.-.-).
GN MACAB OR RSC1088 OR RS04099.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
OC Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=GM11000;
RX MEDLINE-21681879; PubMed-11823852;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangelot S.,
RA Ariat M., Billault A., Brothier P., Camus J.C., Cattolico L.,
RA Chaudier M., Choise N., Claudel-Renard C., Cunnac S., Demange N.,
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,
RA Sigulier P., Thebaud P., Whalen M., Wincker P., Levy M.,
RA Weissenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum."
RL Nature 415:497-502(2002).
DR EMBL; AL646062; CAD14790.1; -
DR InterPro; IPR001281; Rieseke.
DR Pfam; PF00355; Rieseke; 1.
KW Oxidoreductase; Complete proteome.
SO SEQUENCE 103 AA; 11414 MW; 81CBB708912AED2A CRC64;

Query Match
Best Local Similarity 100.0%; Score 21; DB 16; Length 103;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HARL 4
DB 49 HARL 52

RESULT 45
Q9D6F0 PRELIMINARY; PRT; 104 AA.
ID Q9D6F0;
AC Q9D6F0;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE 2900064B16RIK protein.
GN 2900064B16RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=HIPPOCAMPUS;
RX MEDLINE-21085660; PubMed-11217851;
RA Kawai T., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Araiawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K. I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischnann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staudli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Baird G.,

```

RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyono-Oka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
 RA Wyszynski-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
 RA Hayashizaki Y.,
 RT "Functional annotation of a full-length mouse cDNA collection.",
 RL Nature 409:685-690(2001).
 DR EMBL; AK013746; BAB28981.1; -;
 DR MGD; MGI:1920284; 2900064B16Rik.
 SO SEQUENCE 104 AA; 11669 MW; 7BAD8049EC63D6B0 CRC64;

Query Match 100.0%; Score 21; DB 11; Length 104;
 Best Local Similarity 100.0%; Pred. No. 5.9e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
 ||||
 DB 96 HARL 99

RESULT 46
 ID 080ZYS PRELIMINARY; PRT; 105 AA.
 AC 080ZYS;
 DT 01-JUN-2002 (TREMBLrel. 21, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Fus 2 (Fragment).
 GN FUS2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C129/SV;
 RA Shuttleworth T.L., Wilson M.D., Wicklow B.A., Wilkins J.,
 RA Triggs-Raine B.L.,
 RT "Characterization of the Murine Hyaluronidase Gene Region Reveals
 RT Complex Organization and Co-transcription of Hyal1 with Downstream
 RT Genes, Fuz2 and Hyal3.".
 RL J. Biol. Chem. 0:0-0(2002).
 DR EMBL; AF417493; AAM14423.1; -;
 DR EMBL; AF417497; AAM14431.1; -;
 FT NON_TER 105 105
 SO SEQUENCE 105 AA; 11522 MW; 859753EE6033F919 CRC64;

Query Match 100.0%; Score 21; DB 11; Length 105;
 Best Local Similarity 100.0%; Pred. No. 5.9e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
 ||||
 DB 52 HARL 55

RESULT 47
 ID 090TE6 PRELIMINARY; PRT; 106 AA.
 AC 090TE6;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update).
 DE Homolog of MDV1 pp24 protein (Fragment).
 OS Marek's disease virus serotype 2 MDV2.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Marek's disease-like viruses.
 OX NCBI_Taxid=36353;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=HPRS24;
 RA Jang H., Cai J., Izumiya Y., Murakami Y., Mochizuki M., Song C.,
 RA Lee I., Kai C., Takahashi E., Mikami T.,
 RT "The complete DNA sequence and transcription map of the unique long
 RT genome region of Marek's disease virus type 2.".
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB024414; BAA82891.1; -;
 FT NON_TER 1 1
 SO SEQUENCE 106 AA; 11066 MW; 275DF75DF3E15C39 CRC64;

Query Match 100.0%; Score 21; DB 12; Length 106;
 Best Local Similarity 100.0%; Pred. No. 6e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
 ||||
 DB 34 HARL 37

RESULT 48
 ID 09F8A0 PRELIMINARY; PRT; 107 AA.
 AC 09F8A0;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE MOBC.
 GN MOBC.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia
 OX NCBI_Taxid=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LG74;
 RA Zakharova M.V.,
 RT "Escherichia coli plasmid pLG13, complete sequence.".
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF251289; AAG18127.1; -;
 KW Plasmid.
 SO SEQUENCE 107 AA; 12019 MW; 15DE455470D720C7 CRC64;

Query Match 100.0%; Score 21; DB 2; Length 107;
 Best Local Similarity 100.0%; Pred. No. 6e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HARL 4
 ||||
 DB 11 HARL 14

RESULT 49
 ID 069423 PRELIMINARY; PRT; 107 AA.
 AC 069423;
 DT 01-AUG-1998 (TREMBLrel. 07, Created)
 DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE MbEC.
 GN MbEC.
 OS Enterobacter cloacae.
 OC Bacterioidetes; Bacterioidetes; gamma subdivision; Enterobacteriaceae;
 OC Enterobacter.
 OX NCBI_Taxid=550;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=18K;
 RA Solonin A.S.,
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; Y16897; CAA76522.1; -;

KW Plasmid.
SQ SEQUENCE 107 AA; 11886 MW; 1372246ACF372182 CRC64;

Query Match 100.0%; Score 21; DB 2; Length 107;
Best Local Similarity 100.0%; Pred. No. 6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HARL 4
1111
DB 11 HARL 14

RESULT 50

Q52252 PRELIMINARY; PRT; 107 AA.
ID Q52252
AC Q52252;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE Mobilization protein.
GN MBEC.
OS Salmonella enterica.
OG Plasmid pMO799.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=28901;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN-BORREZE;
RX MEDLINE=95394842; PubMed=7545154;
RA Keenleyside W.J., Whitfield C.;
RT "Lateral transfer of rfb genes: a mobilizable COLE1-type plasmid carries the rfbO:54 (O:54 antigen biosynthesis) gene cluster from RT Salmonella enterica serovar Borreze.";
RL J. Bacteriol. 177:5247-5253(1995).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN-BORREZE;
RX MEDLINE=97067086;
RA Keenleyside W.J., Whitfield C.;
RT "A novel pathway for O-polysaccharide biosynthesis in Salmonella RT enterica serovar Borreze.";
RL J. Biol. Chem. 271:28581-28592(1996).
[3]
RP SEQUENCE FROM N.A.
RC STRAIN-BORREZE;
RX MEDLINE=97158136; PubMed=9004408;
RA Reeves P.R., Hobbs M., Valvano M.A., Skurnik M., Whitfield C., Coplin D., Kido N., Klena J., Maskell D., Raetz C.R., Rick P.D.;
RT "Bacterial polysaccharide synthesis and gene nomenclature.";
RL Trends Microbiol. 4:495-503(1996).
DR EMBL: L39794; AAC98397.1; -;
KW Plasmid.
SQ SEQUENCE 107 AA; 11952 MW; 13722946CA737F7 CRC64;

Query Match 100.0%; Score 21; DB 2; Length 107;
Best Local Similarity 100.0%; Pred. No. 6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HARL 4
1111
DB 11 HARL 14

Search completed: November 13, 2002, 13:42:15
Job time : 31 secs